The influence of different forms of outcome information on the neural substrates of the acquisition and representation of categories

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

genehmigt durch die Fakultät für Naturwissenschaften
der Otto-von-Guericke-Universität Magdeburg

von Dipl. Psych. Reka Daniel-Weiner (geb. Daniel)
geb. am 20. November 1982 in Veszprem

Gutachter: Prof. Dr. Stefan Pollmann
Prof. Dr. Christiane Thiel

eingereicht am: 20. April 2012
verteidigt am: 17. Juli 2012

Abstract

The ability to divide objects and events into separate categories based on their functional relevance is a central determinant of both behavior and cognition (Ashby & Maddox, 2005; Harnad, 2005). Complex category structures are known to be acquired best in the presence of response-contingent feedback. This effect is supposed to be mediated by the mesencephalic dopaminergic system (Ashby & Maddox, 2005), which has been studied mainly in the context of reward-based learning (Niv & Montague, 2008). However, in everyday learning most decisions are not immediately followed by reward or external information on correctness (Hogarth, 2006). The set of experiments presented in this thesis was designed to test the influence of the nature of outcome information on the neural substrates of the acquisition of categories and their representation. Three types of outcome information were compared: monetary reward, cognitive feedback and internally generated signals on correctness. All experiments were carried out using functional magnetic resonance imaging (fMRI) on healthy young human participants.

In the first experiment, two parallel category learning tasks were developed. Participants received a monetary reward for correct answers in one task, in the other task only cognitive feedback on correctness was provided. A post-experimental questionnaire assessed the motivational state during categorization and fMRI activations in the two tasks were compared (Daniel & Pollmann, 2010). This approach was extended in a second experiment. Participants learned the task via observation without ever receiving feedback and were questioned about their confidence after each decision as a measure of internal feedback signals. The neural correlates of these internal signals were examined in terms of reinforcement learning theory (Daniel & Pollmann, 2012). In a third experiment, the lasting effects of monetary reinforcement and feedback on correctness during training were compared. Here, similarity patterns of fMRI activation during passive viewing after successful acquisition of the category structure were assessed.

Results show a parallel implication of many major target structures of the dopaminergic system, including the ventral striatum, during category learning with explicit rewards, cognitive feedback, and in response to internal signals on correctness. Yet, both qualitative and quantitative differences in activation are reported. During category learning effects in the ventral striatum are smaller in the absence of external rewards, and in visual areas the representation of the categorical structure of stimuli is less pronounced after feedback-based than after reward-based training. Also, the motivational states induced by reward influence the pattern of activation, and a specific representation of potentially more valuable stimuli in the orbitofrontal cortex can be observed.

The experiments presented in this thesis provide a parsimonious account of learning in ecologically valid settings, where outcome information does not always take the form of explicit rewards, by showing that structures of the dopaminergic system are activated irrespective of the nature of feedback. Additionally, several differences in implication are pointed out, which help to clarify the neural basis of previously observed differences due to the specific training protocol.

Acknowledgements

Firstly, I want to thank Stefan Pollmann, who, as my supervisor, conceived the idea for this project, but also gave me a lot of room to develop my own take on it. Despite the freedoms I enjoyed working in his lab, he was always available when I had questions and needed input. Most importantly, he knew when to intervene and stop me from adding just another control condition or calculating a further analysis, and to get the paper (and this thesis) finally out there.

I am very grateful to all the people within Magdeburg who made the SFB 779, a collaborative research center on the neurobiology of motivated behavior, possible. Not only did it fund this project, but it also allowed me to have a close look at research on motivation from many different angles, from the protein to the whole human brain.

Jana Tegelbeckers, who accompanied this project as a student assistant from its conception to the end, always seemed to have a trick up her sleeve when in the last minute participants became unavailable, and her organization of data acquisition was unnoticeable since everything worked so smoothly. I thank Jana for the extra hours she put into this project and wish her all the best with her own dissertation, which she recently started within the SFB 779.

I also want to thank the participants in my experiments; I am aware that some of the tasks were less than fun to perform. Without people who are willing to invest time and effort to provide data, empirical research on humans would be quite pointless. Also, I am grateful to the anonymous reviewers who offered their comments on the already published parts of this thesis during the peer review process – these comments have substantially improved the present work.

Without my colleagues the last years would have been both a lot less fun and a lot less fruitful. Florian Baumgartner, Franziska Geringswald, Michael Hanke, Angela Manginelli, Sascha Purmann and, in the last months, Barbara Wolynski and Wolf Zinke were not only always available for theoretical and methodological discussions, but also for sharing meals (in certain periods every meal of the day), coffee breaks, and leisure time.

Lastly, I want to thank my husband Johannes Weiner, who supported and encouraged me to move to Magdeburg for this project, and joined me there for the last year. He managed to brighten my mood whenever things did not go quite as planned, and reminded me that there is a life outside the (home) office.

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA . . . . . analysis of variance

BA . . . . . Brodmann area

BIC . . . . . Bayesian information criterion

BOLD . . . . . blood oxygenation level-dependent

COVIS . . . . . COmpetion between Verbal and Implicit Systems

CS . . . . . conditioned stimulus

dACC . . . . . dorsal anterior cingulate cortex

dmPFC . . . . . dorsomedial prefrontal cortex

EEG . . . . . electroencephalography

EPI . . . . . echo planar imaging

ERN . . . . . error-related negativity

FA . . . . . flip angle

fMRI . . . . . functional magnetic resonance imaging

FoV . . . . . field of view

FRN . . . . . feedback-related negativity

FWE . . . . . family-wise errorrate

FWHM . . . . . full width at half maximum

GLM . . . . . general linear model

HRF . . . . . hemodynamic response function

HWHM . . . . . half width at half maximum

IMI . . . . . Intrinsic Motivation Inventory

ITC . . . . . inferior temporal cortex

ITI . . . . . intertrial interval

lPFC . . . . . lateral prefrontal cortex

MNI . . . . . Montreal Neurological Institute

MP-RAGE . . . . . magnetization-prepared rapid acquisition gradient echo

mPFC . . . . . medial prefrontal cortex

MRI . . . . . magnetic resonance imaging

MTL . . . . . medial temporal lobe

OFC . . . . . orbitofrontal cortex

PET . . . . . positron emission tomography

PFC . . . . . prefrontal cortex

RCZ . . . . . rostral cingulate zone

ROI . . . . . region of interest

SD . . . . . standard deviation

SN . . . . . substantia nigra

TE . . . . . time to echo

TI . . . . . inversion time

TR . . . . . time to repeat

US . . . . . unconditioned stimulus

vmPFC . . . . . ventromedial prefrontal cortex

VTA . . . . . ventral tegmental area

1 General Introduction

1.1. Motivation and structure of this thesis

Learning which action leads to the most beneficial outcome in a given situation is one of the central components of adaptive behavior, and choices are often guided by experience about which behavior provided the most favorable outcome in similar situations in the past. In the last decade, impressive advances in examining the neural bases of reinforcement learning have been made. Computational theories suggest that during reinforcement learning associative links between stimuli and outcomes are formed and adjusted after each trial to minimize future errors in prediction of the outcome (Rescorla & Wagner, 1972; Sutton, 1988). Although the first observations on the relevance of structures of the dopaminergic system for reward-related processes have been performed in rodents (Olds & Milner, 1954), its importance for reward-based learning in humans has received considerable support from studies using fMRI. It has been shown to code variables from normative computational reinforcement learning models like reward expectation and the reward prediction error in the human brain during decision making. Especially the ventral striatum is activated by a wide range of reinforcing stimuli including juice, odors, money and beauty (Abler, Walter, Erk, Kammerer, & Spitzer, 2006; Aharon et al., 2001; Gottfried, O'Doherty, & Dolan, 2002; McClure, Berns, & Montague, 2003).

Many of these studies have employed instrumental conditioning tasks. Typically, a limited set of distinct stimuli is presented to the participant, and after each response immediate reward is available. However, in ecologically valid settings organisms are often faced with more complex decisions problems. They have to extract the current state of the environment from the noisy sensory signals they receive, and to interpret these signals in terms of their functional relevance. A fundamental component of this process involves categorization. Ashby and Maddox (1998) define a categorization task as one where there are more stimuli than responses, which implies that basic categorizations serve to map the wide and often continuous range of sensory stimulation onto a limited number of behavioral responses. Complex category learning tasks, which force participants to rely on gradually acquired stimulus-outcome contingencies, are sensitive to the nature and timing of feedback (Maddox, Ashby, & Bohil, 2003; Maddox, Love, & Glass, 2008). Yet, in everyday learning, after many decisions immediate reward is not available. Often only information about the correctness of the decision is provided, and in other cases even this information has to be inferred from previously acquired knowledge (Hogarth, 2006).

The three experiments presented in this thesis aim at examining the influence of the exact nature of outcome information on the neural correlates underlying the acquisition and representation of categories. To this end, the activations in response to internally generated signals on correctness, cognitive feedback, and monetary reward during and after categorization training will be compared using fMRI. In the remainder of this first chapter I will review current theories and empirical results on category learning, on the ability of fMRI to assess neural and especially dopaminergic activations, and on reinforcement learning. Also, the open questions addressed by this thesis will be summarized. Together with Chapter

2, which introduces common methods of all presented experimental work, this serves to provide the background for the experiments reported in Chapters 3 to 5. Finally, the results are summarized and their implications are discussed in Chapter 6.

1.2. The cognitive neuroscience of human category learning

The ability to divide objects and events into separate, meaningful categories is both a basic skill of any living organism and a fundamental characteristic of sophisticated thought (Seger & Miller, 2010; Ashby & Ennis, 2006). The central importance of categorization to organized behavior is most evident when imagining an organism that lacks this ability: sensory inputs will be experienced as unfamiliar as soon as they differ only slightly from the inputs during previous events and the world will be experienced as a stream of fragmented and unrelated events, a „blooming, buzzing confusion" (James, 1890). Most importantly in the context of cognition, the lack of the ability to detect the commonalities between experiences and their higher-level structure prevents such an organism from assigning meaning to sensory experiences, to recognize and respond to objects never encountered in this specific form before, and to anticipate and generalize to future situations by implying that they share basic elements with past experience (Seger & Miller, 2010). This prohibits learning from previous encounters of similar situations, as only after objects and events have been dissociated into meaningful entities can they be associated with different outcomes, which is one of the most basic forms of learning. Because of this fundamental importance, categorization has even been suggested to *be* cognition (Harnad, 2005). Following this line of argumentation it is not surprising that the exploration of category learning has a large and rich history in cognitive sciences, and that a thorough review of this complex phenomenon is well beyond the scope of this thesis. For an introduction from different areas of cognitive science, including psychology, linguistics, philosophy, anthropology, and neuroscience, the reader is referred to H. Cohen and Lefebvre (2005).

The following short overview introduces human perceptual category learning viewed from the perspective of cognitive neuroscience. This focus excludes those parts of the literature on animal categorization that are not considered directly relevant for drawing inferences about human categorization, formal theories of categorization behavior without specification of its underlying substrates, and the literature on concepts. Concepts have been defined as a group of related ideas as opposed to a perceptual category (Ashby & Maddox, 2005), which is defined here as collection of perceptual objects sharing the same functional relevance. Also, as the focus of this introduction is on learning, literature on innate categorization behavior, e.g. in the context of color or speech perception, and the categorization behavior of highly trained experts, are not reviewed. Recent reviews on the latter topic are available elsewhere (Mahon, 2009; Martin, 2007).

1.2.1. The neural substrates of category learning

With the development of new non-invasive methods to monitor the activity of the working brain, in the last decade the focus of research on category learning has shifted from an abstract description of the cognitive processes mediating it, to its neural substrates. To study category learning, i.e. the acquisition of formerly unfamiliar categories, experimentally, it is important to construct new artificial categories to ensure that the participants have not encountered them before. Several methods exist to construct such categories, and depending on the construction method as well as on the method and duration of training, differences in the neural substrates subserving category learning have been observed. In the following sections, some the most common categorization paradigms will be introduced along with the current findings on their neural substrates. The aim of this introduction is threefold: it is intended to provide the necessary background to appreciate the rationale behind the construction of the categorization tasks reported in Chapters 3-5, to introduce the brain network supposed to underlie category learning, and to alert the reader to the limited scope of conclusions which can be drawn from a single categorization

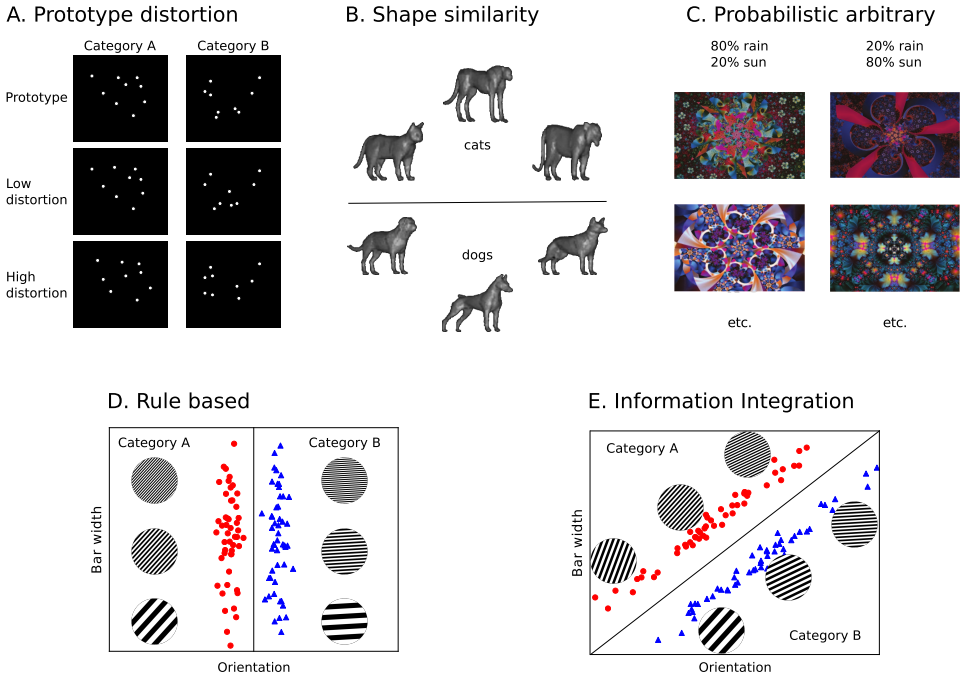

Figure 1.1. *Common category learning tasks.* **A.** Dot pattern prototype distortion task. A prototypical stimulus is created by randomly distributing nine dots, and category exemplars are created by shifting each dot according to a statistical rule. Depending on this rule the amount of dislocation can be varied, resulting in high or low distortions (adapted from Daniel (2008)). **B.** The "cat"-"dog" categorization task from Freedman, Riesenhuber, Poggio, and Miller (2003). Three prototypes from the "cat" and "dog" categories are shown; each prototype can be morphed into every other prototype to create exemplars. **C.** Weather prediction categorization task, adapted from Lopez-Paniagua and Seger (2011). Stimuli within a category do not share common identifying features, and no classification rule can be derived. Each stimulus is probabilistically associated with the categories "rain" or "sun". **D.** + **E.** Rule based and information-integration tasks. Each red circle denotes the coordinates of a stimulus from Category A, each blue triangle denotes the coordinates of a stimulus from Category B. The square-wave gratings are example stimuli constructed from the coordinates. A verbal rule for classification in **D.** might be "All stimuli tilted upwards from "2" on a clock are from category A, all others are from category B". No easily verbalizable rule can describe the category boundary in **E**.

paradigm.

Sensory areas

As the experiments presented in this thesis and the majority of previous studies have been conducted using visual stimulation, I will focus on this domain to discuss the potential contribution of areas involved in sensory processing. A very basic example of a visual perceptual category learning task is the prototype distortion task (Figure 1.1 A). In such tasks exemplars are created by randomly distorting the prototype which defines the category. In the (A, not A)-version participants learn about category membership by passively observing exemplars of a single category. During the task, they have to decide whether presented stimuli are members of the previously presented category, or are a random pattern. It has been suggested that for such simple tasks the locus of plasticity is in early visual areas, mainly V2 (Aizenstein et al., 2000; Reber, Stark, & Squire, 1998b, 1998a), and that activation differences in those early areas are sufficient for making a categorization response (Ashby & Maddox, 2005; E. E. Smith, 2008). However, already in slightly more complex versions of this task, e.g. when participants are required to distinguish between two different categories, and when learning via trial and error, a far more widespread network

including prefrontal, parietal and subcortical areas is thought to subserve prototype distortion learning (Boettiger & D'Esposito, 2005; Daniel et al., 2011; Seger et al., 2000; Vogels, Sary, Dupont, & Orban, 2002).

Neurons at higher levels of the ventral visual processing stream, i.e. in the lateral occipital and inferior temporal cortex, are known to respond selectively to more complex shapes, e.g. to faces (Desimone, Albright, Gross, & Bruce, 1984; Kanwisher, McDermott, & Chun, 1997) or trees (Vogels, 1999). Task independent category information can be extracted from the inferior temporal cortex (ITC) (Kriegeskorte, Mur, Ruff, et al., 2008), and during the categorization of complex shapes neurons in this area are known to emphasize diagnostic features (Sigala & Logothetis, 2002), show greater activity to stimuli near the category boundary (DeGutis & D'Esposito, 2007), and to show a sharper tuning of responses to stimuli across category boundaries after categorization training (Freedman et al., 2003) (for an example of a categorization task employing complex shapes see Figure 1.1 B). As, however, neurons in the ITC do not completely generalize across category members (Eger, Ashburner, Haynes, Dolan, & Rees, 2008; Freedman et al., 2003; Jiang et al., 2007), the ITC has been suggested to contribute to categorization mainly by providing a higher level analysis of the object features relevant for classification (Freedman, 2008; Seger & Miller, 2010).

Prefrontal and parietal cortices

After extensive training the neural firing patterns within the lateral prefrontal cortex (lPFC) reflect the relevant category boundary during the categorization of complex shapes, showing a sharp tuning to category boundaries (Freedman, 2008). A similar activation pattern has been observed in posterior parietal neurons during the classification of moving dot patterns (Freedman & Assad, 2006). Modulated neuronal activity in the lPFC during categorization was has also been shown in humans (Moore, Cohen, & Ranganath, 2006; Jiang et al., 2007), as well as increasing activation with categorization uncertainty in dorsal medial prefrontal areas (Grinband, Hirsch, & Ferrera, 2006; Daniel et al., 2011) and, in a complex dot-pattern classification task, in the posterior parietal cortex (Daniel et al., 2011). These activation patterns are well in line with the general role of the parietal and prefrontal cortices in executive processing, including visuospatial attention and generation of context-appropriate responses, working memory, conflict monitoring and the cognitive control of ongoing behavior (Corbetta & Shulman, 2002; Koechlin, Ody, & Kouneiher, 2003; Miller & Cohen, 2001; Ridderinkhof, Wildenberg, Segalowitz, & Carter, 2004). It is unclear whether the mentioned prefrontal and parietal areas respond specifically to categorization as compared to other tasks with similar processing demands (Grinband et al., 2006).

Supplementary motor and motor cortices

Neural processes involved in the preparation of a potential action may play a general role in the formation of perceptual decisions (Gold & Shadlen, 2001, 2003). Electrophysiological results suggest that sensory signals, which can be used to select the motor response, are conveyed to the primary motor and premotor cortex (Romo, Hernández, & Zainos, 2004; Shen & Alexander, 1997; Zhang, Riehle, Requin, & Kornblum, 1997), and fMRI data recorded from human participants shows that supplementary and frontal eyefields participate in learning in prototype distortion tasks (Little & Thulborn, 2005). The prefrontal cortex (PFC) might be especially important in early category learning, when flexible sets of rules are maintained and tested, while later in training motor areas might gain importance (Seger & Miller, 2010). In accordance with this suggestion, an increasingly more important role of motor-related areas in visual object matching and categorization has been observed as training progressed (Helie, Roeder, & Ashby, 2010; Pollmann & Maertens, 2005; Waldschmidt & Ashby, 2011).

Medial temporal lobe and hippocampus

The hippocampus and surrounding medial temporal lobe (MTL) structures are involved in storing unique experiences and in the rapid learning of individual instances (McClelland, McNaughton, & O'Reilly,

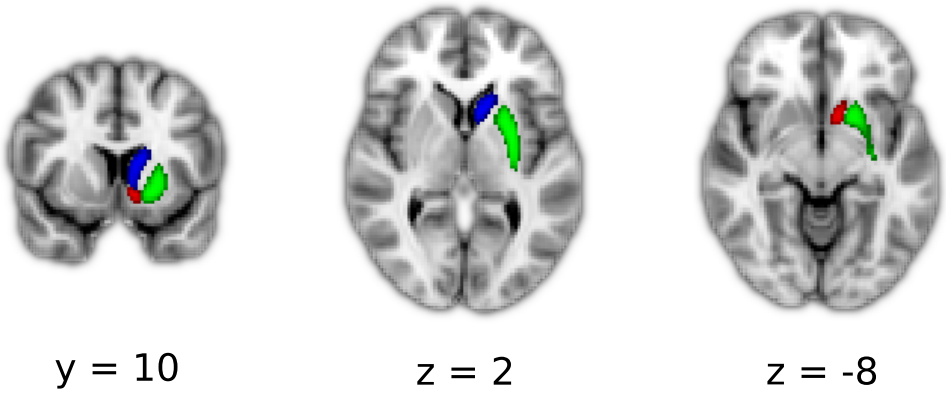

Figure 1.2. *The striatum.* Nucleus accumbens (red), caudate nucleus (blue), and putamen (green). Colors added for one hemisphere based on the Harvard-Oxford subcortical structural atlas implemented in the FSL-FMRIB Software Library (http://www.fmrib.ox.ac.uk).

1995). In the context of category learning they might be required for the initial encoding of novel stimuli and novel associations between stimuli and the category (Meeter, Radics, Myers, Gluck, & Hopkins, 2008). This is especially important in the arbitrary categorization task (see Figure 1.1 C), in which stimuli within a category do not share common features and therefore have to be remembered individually (Poldrack et al., 2001; Seger & Cincotta, 2005). The MTL and hippocampus might however also be involved in other categorization tasks, e.g. in keeping track of rules have that already been tested and rejected (Ashby & Valentin, 2005), in storing exceptions to category-defining regularities (Love, Medin, & Gureckis, 2004), or in generalizing and transferring knowledge (Shohamy, Myers, Kalanithi, & Gluck, 2008; Shohamy & Wagner, 2008).

Basal ganglia and corticostriatal loops

The basal ganglia are a group of subcortical nuclei which have initially been considered to play a central role in motor or sensory-motor function, but since then have been shown to also participate in a wide range of goal-directed behaviors, including motivation and cognition (Haber & Knutson, 2009). They receive projections from virtually the whole cortex via their input structure, the striatum (consisting of the putamen, the caudate nucleus and the nucleus accumbens, see Figure 1.2). These inputs are characterized by a massive convergence of approximately 10,000 to 1 (Ashby & Ennis, 2006) and feed activation back to the cortex via a thalamic pathway. The connections from the cortex and back again maintain some degree of topographic separation and therefore have been suggested to be organized in four interacting corticostriatal loops (see Figure 1.3): the *motor loop* connecting the motor cortex with the posterior putamen, the *executive loop* connecting the PFC and the parietal cortex with the anterior caudate nucleus (head), the *motivational* loop connecting the ventral striatum with the orbitofrontal cortex, and the *visual loop* connecting extrastriate and inferotemporal cortices with the posterior caudate nucleus (body and tail) (Seger, 2006, 2008; Seger & Miller, 2010). In accordance with this unique position in the brain, the basal ganglia, and especially the striatum, participate in a wide range of categorization tasks and are known to play various important roles in category learning. Seger and Miller (2010) suggest that the visual loop feeds information from the visual cortex forward to the executive and motor loops, and back to the visual cortex to modulate visual processing. The executive loop is involved in feedback processing, working memory updating, and set shifting, while the motor loop is associated with selecting and executing the response. Lastly the motivational loop is suggested to be involved in the processing of reward and feedback, as well as in integrating this information across time for future responses (Seger, Peterson, & Cincotta, 2010; Lopez-Paniagua & Seger, 2011).

Early research on the neural substrates of category learning has shown that the presence of feedback is a

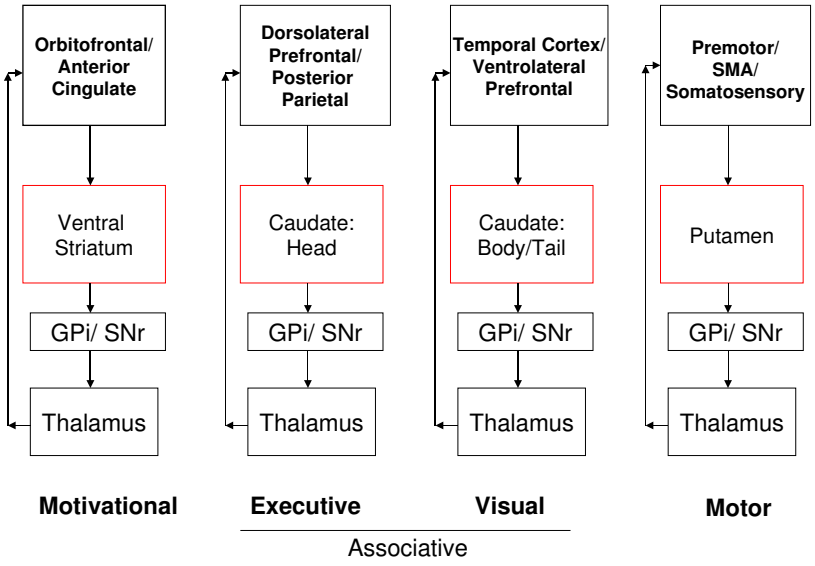

Figure 1.3. *Corticostriatal loops.* GPi: Globus pallidus, internal portion, SNr: Substantia nigra pars reticulata (based on Seger (2008)).

central determinant for the implication of the striatum in category learning. In an arbitrary classification task, in which the correct classification rule was difficult to describe verbally due to probabilistic stimulus-response-outcome contingencies, Poldrack et al. (2001) showed that activation in the caudate is higher during feedback-based learning than during observational learning. Also, patient populations with impaired striatal functioning are impaired only in feedback-based category learning and mainly in complex categorization tasks which require incremental learning and the integration of information across multiple experiences (Shohamy et al., 2008). However, these results on the feedback-dependence of striatal involvement have to be differentiated for the subcomponents of the striatum: activity in the head of the caudate was observed to be modulated by the presence of feedback, while the posterior caudate and the putamen were active to a similar degree in both feedback-based and observational learning (no data reported on the ventral striatum; Cincotta and Seger (2007)). Also, the ventral striatum is activated during both categorization and feedback processing, but the activation is stronger during classification (Aron et al., 2004; Lopez-Paniagua & Seger, 2011). This indicates that once active in the learning process, ventral striatal activity is not limited to the feedback component of the task (Poldrack & Willingham, 2006).

In addition to the presence or absence of feedback, two further factors are known to influence the respective contributions of the corticostriatal loops: the amount of expertise acquired, and the specific properties of the task. The executive and motivational loops might be most important early in learning, whereas the motor loop becomes more involved when the task is well-learned (Seger & Cincotta, 2005). An empirically well-supported theory of category learning, COmpetion between Verbal and Implicit Systems (COVIS) (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby, Paul, & Maddox, 2011), proposes a competition between the visual and executive loops dependent on the category structure. In the framework of COVIS rule-based category learning tasks (see Figure 1.1 D) are defined as tasks in which the category structure can be learned via explicit reasoning and the rule maximizing accuracy is

easy to describe verbally. In these tasks the verbal system, consisting of the executive loop through the head of the caudate nucleus and additional loops through sensory association areas and the MTL, is assumed to predominate (Ashby & Valentin, 2005). The learning of rule-based tasks has been shown to be insensitive to presence and timing of feedback, which has been attributed to the reliance of the verbal system on working memory and attention (Maddox et al., 2003; Ashby & Maddox, 2002). In contrast, in information-integration category learning (see Figure 1.1 E) accuracy is maximized only if information from two or more stimulus components / dimensions is integrated at a pre-decisional stage and the optimal strategy is difficult or impossible to describe verbally. Here, the implicit system based on the visual loop is assumed to predominate eventually, as it is more efficient in learning complex implicit rules, however does not learn as rapidly as the explicit system (Ashby & Maddox, 2005). The learning of information-integration tasks is postulated to be procedural (Ashby & O'Brien, 2007) and highly dependent on feedback. Information-integration category structures cannot be learned using unsupervised training protocols (Ashby, Queller, & Berretty, 1999), and the acquisition of the category boundary is impaired when training is observational rather than feedback-based (Ashby & Maddox, 2002), or feedback is delayed (Maddox et al., 2003; Maddox & Ing, 2005). These effects are attributed to the dependence of the implicit system on reward-based teaching signals originating in the dopaminergic system (Ashby & O'Brien, 2007). COVIS assumes that both the verbal and implicit system remain active throughout the whole process of category learning. When the task is overlearned, the subcortical pathway through the basal ganglia is suggested to be gradually replaced by a more direct cortico-cortical path from perceptual to motor areas (Ashby & O'Brien, 2007; Helie et al., 2010; Waldschmidt & Ashby, 2011).

Dopaminergic midbrain nuclei

As all forms of reinforcement-based learning, feedback-based category learning critically depends on the function of the dopaminergic midbrain nuclei. The striatum is both a major contributor of input to the dopaminergic system, and a major recipient of dopaminergic output (Joel & Wiener, 2000). Within the striatum, learning is assumed to be dependent on three factors: strong presynaptic activation, strong postsynaptic activation, and dopamine release. This indicates that striatal learning, which is central to many forms of category learning, can only occur when the dopaminergic system is active (Ashby & Ennis, 2006). Neuropsychological data shows that patients with Parkinson's disease, which is characterized by a loss of dopaminergic input to the striatum, are impaired in feedback-based complex category learning tasks, and that this effect is modulated by dopaminergic medication (Ashby & Maddox, 2005; Shohamy et al., 2008). Reinforcement learning theory provides an elaborated account of dopaminergic functioning. As the topic of feedback processing is central to the experiments presented in this thesis, a more thorough overview of reinforcement learning theory is provided in Section 1.4.

1.2.2. Summary and outlook

Not surprisingly for such an ancient and fundamental skill, depending on the specifics of the task virtually all brain areas, including the neocortex (sensory, prefrontal, parietal, and motor cortex), the medial temporal lobe, basal ganglia and dopaminergic midbrain neurons have been shown to subserve category learning.

Three major conclusions can be drawn from the literature reviewed in this section: the neural structures underlying category learning are strongly dependent on (1) the task itself, (2) the level of expertise acquired in the task, and (3) the way the task is trained. For simple classification tasks perceptual learning within early sensory areas might be sufficient, while highly trained tasks might only involve connections from sensory to motor cortices. In contrast, the initial acquisition of more complex category structures relies on corticostriatal loops. Predominantly in tasks in which the optimal categorization rule is difficult to describe verbally, learning is strongly dependent on feedback, and the effects of feedback

are carried by the dopaminergic system. As the focus of the present thesis is on the influence of outcome information, such a category learning task will be employed in all experiments presented.

Most experiments on the involvement of the dopaminergic system in learning in human participants have been conducted using fMRI. Therefore, in Section 1.3 an introduction to the origin of the signal measured using fMRI is provided, along with a critical discussion of its ability to assess dopaminergic activation. In Section 1.4 the most central models of reinforcement learning are presented, and empirical findings on the involvement of the dopaminergic projection areas in human learning are summarized.

1.3. Examining the brain using fMRI: the relationship between the fMRI signal, neuronal firing and dopaminergic activity

FMRI combines a moderate spatial resolution in the range of millimeters with a moderate temporal resolution in the range of seconds. It is able to simultaneously provide data from the whole brain, including subcortical structures, which enables the examination of the entire network of areas engaged while participants are performing a task. In combination with its relative cost-efficiency and ever-increasing availability, these properties have made fMRI one of the most popular methods to measure brain activity in humans. Since its introduction in 1992 (Bandettini, Wong, Hinks, Tikofsky, & Hyde, 1992; Frahm, Bruhn, Merboldt, & Hänicke, 1992; Kwong, 1992; Ogawa, 1992), this method has experienced an exceptional growth. In 2008, a query of the ISI / Web of Science database using the keywords 'fMRI' or 'functional MRI' or 'functional magnetic resonance imaging' returned around 19,000 peer-reviewed articles (Logothetis, 2008). In 2012, the same search returns over 50,000 articles.

Excellent introductory books on the method of fMRI are available (e.g. Ashby (2011), Huettel, Song, and McCarthy (2004), Poldrack, Mumford, and Nichols (2011), and Jezzard, Matthews, and Smith (2001)), and a short overview by the author of this thesis is provided in Daniel (2008)[1]. Given the widespread recognition of fMRI no general introduction to the method itself will be provided here. However, the experiments presented in Chapters 3-5 rely on the fMRI signal to examine the neural correlates of learning and feedback processing. As these processes are intimately tied to the activity of dopaminergic midbrain neurons and their projections, this section provides a short discussion of the origin of the fMRI signal in general and of its relationship to dopaminergic activity in particular.

1.3.1. Understanding the fMRI signal

Despite its wide use, the exact relationship between the fMRI signal and the underlying neuronal activity, which it is assumed to represent, is not yet fully understood. As the fMRI technology cannot provide direct information about the electrical activity of neurons, it has to rely on an indirect index based on the metabolic changes induced by neuronal activity. Hemoglobin, the oxygen-carrying molecule within red blood cells, has different magnetic properties depending on its oxygenation state, which allows it to act as an endogenous contrast agent. The blood oxygenation level-dependent (BOLD) contrast is a compound signal which reflects changes in oxygen consumption in combination with changes in cerebral blood flow and volume (Buxton, Uludağ, Dubowitz, & Liu, 2004). All of these vascular responses are elicited by local increases in neural activity (Logothetis & Wandell, 2004; Logothetis, 2010); however, the coupling mechanisms between the BOLD signal and neural activity are indirect, complex and interrelated. They depend on factors like the type of neural response to a stimulus, the link between this response and energy demands, the mechanisms that signal the energy demand, and the mechanisms responsible for triggering the hemodynamic response (Arthurs & Boniface, 2002; Logothetis, 2003).

Simultaneous cortical electrophysiological and fMRI recordings suggest that the BOLD signal is better explained by local field potentials than neuronal spiking activity (Goense & Logothetis, 2008; Kayser, Kim, Ugurbil, Kim, & König, 2004; Logothetis & Wandell, 2004; Mukamel et al., 2005; Raichle & Mintun,

[1] Unpublished diploma thesis. Available from http://apsy.gse.uni-magdeburg.de/~daniel

2006). Local field potentials are assumed to not reflect local action potentials carried by the output neurons of a region, but rather to index the input to a given cortical area as well as its local intracortical processing, including both excitatory and inhibitory interneuronal activity (Logothetis, 2003; Logothetis & Wandell, 2004). In addition, the information reaching a brain region is evaluated under the overall regulation of cortical dynamics by neuromodulatory signals from a number of neurochemicals, including acetylcholine, norepinephrine, serotonin, and dopamine. The neuromodulatory effects affect large masses of cells, and might induce larger changes in the fMRI signal than function specific processing of sensory signals (Logothetis, 2008).

The indirect nature of the BOLD-fMRI signal has to be kept in mind whenever fMRI activity is interpreted as neuronal activation. This is especially true when fMRI activations are interpreted as reflecting the activity of specific neurotransmitter / neuromodulator systems, like it is often done in studies of reward-based learning, which is intimately tied to the function of the dopaminergic system.

1.3.2. A potential link between BOLD response and dopaminergic activation

As detailed in the previous section, fMRI is not able to directly index changes in dopamine release. However, there are indications that dopaminergic neuron firing could be closely correlated with the BOLD response to afferent input in midbrain areas containing a large proportion of dopaminergic neurons, like the substantia nigra (SN) / ventral tegmental area (VTA), and in regions receiving massive projections from these areas, like the ventral striatum.

Injections of dopamine-releasing agents can increase the BOLD signal in the nucleus accumbens mediated by an increased postsynaptic D1 agonism, and it has been suggested that the signal has the capacity to fluctuate at a second-to-second timescale in response to short bursts of dopaminergic firing (Knutson & Gibbs, 2007). Consistent with this assumption, fMRI studies show activations in the vicinity of the midbrain dopaminergic nuclei and in the striatum that correlate with reward expectation and the prediction error, and thereby exhibit a striking similarity to the pattern of burst firing of dopamine neurons observed in electrophysiological recordings (D'Ardenne, McClure, Nystrom, & Cohen, 2008; M. X. Cohen, 2008; O'Doherty, Hampton, & Kim, 2007; Knutson, Taylor, & Kaufman, 2005; Niv, 2009). These results indicate that the association between the presence of dopamine and BOLD activation in the nucleus accumbens might not only occur at the longer timescales associated with pharmacological manipulations, but also at shorter timescales in response to endogenous bursts of dopamine release (Knutson & Gibbs, 2007). More direct evidence for the modulation of ventral striatal fMRI signals by dopamine was provided by Pessiglione, Seymour, Flandin, Dolan, and Frith (2006). They report that during instrumental learning the magnitude of the reward prediction error as measured using BOLD-fMRI in the striatum is modulated by the administration of drugs that enhance or reduce dopaminergic function. Subsequently, a quantitative role of dopamine in the fMRI signal was established by showing a direct correlation between reward-related dopamine release, as indexed by [^{11}C]raclopride positron emission tomography (PET), and fMRI activity in the SN / VTA and the ventral striatum across the same participants (Schott et al., 2008). An indication for the role of dopamine in the dorsal striatum was provided by a recent study by Schönberg et al. (2010). They showed that in patients with Parkinson's disease, which is characterized by a loss of dopaminergic input to mainly the dorsal striatum, fMRI activation in response to reward prediction errors in the dorsal putamen is significantly decreased compared to responses observed in control participants.

Although the studies reviewed above provide evidence for a link between dopaminergic neurotransmission and fMRI activation during reward-learning, they do not rule out that this link is mediated by non-dopaminergic processes, like glutamatergic costimulation. In addition, this evidence was only provided for the expectation and processing of rewards in the SN / VTA and striatum. As fMRI cannot isolate dopamine-related activity from the effects of other afferents, it is conceivable that the BOLD signal and dopamine release dissociate in other situations. Most importantly, the response to aversive stimuli or omissions of expected rewards is associated with increased afferent input to inhibitory interneurons in the SN / VTA (Düzel et al., 2009), which might result in an increased fMRI signal. Indeed, positive midbrain

activations in response to negative feedback and positive errors in the prediction of punishment have been observed (Aron et al., 2004; Menon et al., 2007). The results on a possible relationship between dopaminergic activity and the fMRI signal can also not be transferred to other target areas of the mesencephalic dopaminergic system, like the orbitofrontal cortex or the medial prefrontal cortex, which are less physiologically homogenous and show different pharmacokinetics (Knutson & Gibbs, 2007).

1.3.3. Summary and conclusions

FMRI cannot assess the exact neuronal mechanisms underlying the studied tasks. The signal changes it measures do however consistently reflect local changes in overall neural activity, and it allows measuring these changes with a relatively high spatiotemporal resolution throughout the whole brain, including deep structures that are central to reinforcement learning like the dopaminergic midbrain nuclei and the ventral striatum. For these structures evidence exists that links fMRI activation during reward processing directly to dopamine. A connection with the dopaminergic system has to be assumed as the most probable source of fMRI activations in response to reward in the SN / VTA and nucleus accumbens for the generation of further hypotheses testable with other, more invasive, methods. In combination with its non-invasive nature, these findings make fMRI the best currently available tool to study the neural basis of reward-based learning in healthy human participants.

1.4. Neurocomputational mechanisms of reinforcement learning

As any living organism, humans are faced with the need to make decisions about how to act in response to environmental cues every day. Often, we encounter similar decision problems – or decision problems which we *categorize* to be similar – repeatedly, which gives us the opportunity to learn from previous experience. In behavioral psychology, two classes of conditioning paradigms are distinguished. One class is termed Pavlovian or classical conditioning, and focuses on behavioral changes in response to contingencies between stimuli. As in these paradigms subjects only experience the relationships between events in the world, the changes in behavior are assumed to reflect innately specified reactions to the prediction of the outcomes. The outcomes occur regardless of the subjects' actions, therefore this form of learning sometimes is also referred to and modeled as *prediction learning* (Dayan & Balleine, 2002). The second class is termed instrumental or operant conditioning. In instrumental conditioning paradigms subjects learn to select actions that maximize the outcome of their behavior. Reinforcements are outcomes that increase the probability of the behavior that lead to the outcome, while punishments lead to a decrease in the probability of that behavior. If a behavior is reinforced by providing a stimulus rather than withdrawing it, i.e. by positive reinforcement, the reinforcer is also referred to as reward (Wise, 1989). As in instrumental conditioning the outcome is contingent upon behavior, the subject has the opportunity to maximize this outcome. Therefore, instrumental conditioning can be viewed as the most fundamental form of rational decision-making (Niv, 2009), and is closely related to the computer science theory of reinforcement learning (Sutton & Barto, 1990) and the engineering theory of optimal control (Bertsekas & Tsitsiklis, 1996). These two lines of research study how systems can choose their actions to maximize rewards and minimize punishments, i.e. they provide normative accounts how to *optimize* behavior. Recent research combining normative computational modeling with neuroscientific methods indicates that normative models provide a useful framework for studying reward-based learning in living organisms. A wealth of reinforcement learning models exists which are specifically designed to account for certain phenomena under certain circumstances. In the following section I give a short historical overview of computational models of reinforcement learning as far as they are relevant to the studies that provide the background for the experiments reported in this thesis. Subsequently, evidence is reviewed that these abstract mathematical formulations, which describe how decisions could theoretically be optimized, correlate with neural activations observed in living organisms. A more thorough overview of these topics can be found in Niv (2009) and M. X. Cohen (2008). Detailed descriptions and comparisons

of different reinforcement learning models are provided by Sutton and Barto (1998).

1.4.1. Introduction to computational models of reinforcement learning

The reinforcement learning models introduced in the next sections share some core features. They all calculate a numerical representation for the preference of particular stimuli or actions. Depending on the model these numerical representations are called association strengths, weights, or values. When new information becomes available, these preferences are updated according to an updating term, which is, in its most commonly used form, the reward prediction error, i.e. the difference between received and expected rewards. When faced with a decision problem, the agent selects the stimulus or action with the highest preference value, often after passing the preference values through a probabilistic function (e.g. the *soft-max* distribution) (M. X. Cohen, 2008).

Models of prediction learning

One of the most influential models of classic conditioning is the Rescorla-Wagner model (Rescorla & Wagner, 1972). It distinguishes between the unconditioned stimulus (US), a biologically relevant stimulus that naturally evokes an unconditional (or unconditioned) reaction, and the conditioned stimulus (CS), usually a previously neutral stimulus that comes to evoke a conditioned response through its pairing with the US. The model assumes that after each conditioning trial the associative strength V of the presented conditioned stimuli CS_i changes according to the rule

$$\Delta V_{CS_i} = \alpha_i \beta (\lambda_1 - \sum_i V_{CS_i}) \tag{1.1}$$

where α_i and β are learning rate parameters that depend, respectively, on the CS and the US, and λ_i is the asymptote of learning, i.e. the maximum conditioning possible for the presented unconditioned stimulus. According to this model, if the total of all associative strengths of the conditioned stimuli presented in one trial $\sum_i V_{CS_i}$ is equal to the maximum possible conditioning, associative strengths do not change after a trial. Thereby, the Rescorla-Wagner model postulates that learning only occurs when events violate expectations. Using this error-correcting approach the Rescorla-Wagner model has been successful in explaining and predicting a number of experimental observations like blocking, i.e. the phenomenon that predicted stimuli will not support the conditioning of an additional CS (Kamin, 1969), or of overexpectation, observable as a decline of associative strengths when a pair of two previously well-trained CS is further trained with both CS in compound (Rescorla, 1970).

Based on the Rescorla-Wagner rule, an adaptive network model of category learning was suggested (Gluck & Bower, 1988b, 1988a). In its simplest form this model assumes that a layer of distinct input units – one per presented stimulus – feeds activation directly into the output units representing the categories.

If there are n input nodes, and the activation in input node i is represented by a_i, then the activity in output node o_j, is determined by

$$o_j = \sum_{i=1}^{n} w_{ij} a_i \tag{1.2}$$

and the weights w from input node i to output node j are adjusted according to the error correction rule

$$\Delta w_{ij} = \beta (d_j - o_j) a_i \tag{1.3}$$

where d_j is the desired output for o_j given the input. It can be interpreted as "teaching signal" indicating what the activation of that node should be to obtain the correct response. Even in this parsimonious form, assuming only one layer, this adaptive network model is successful in predicting

both categorization behavior (Gluck & Bower, 1988b, 1988a) and neural activation (Rodriguez, Aron, & Poldrack, 2006) in probabilistic category learning tasks. It is utilized in Chapter 5 to develop an adaptive training method ensuring successful category learning in all participants.

Despite its success, the Rescorla-Wagner model has some serious limitations, one of them being that it explains conditioning on an artificially defined trial-by-trial basis. To overcome this problem in temporal-difference learning models the timepoint t within a trial is explicitly represented (Sutton & Barto, 1990). For a single stimulus S its value V is updated according to the rule

$$V(S_t)_{new} = V(S_t)_{old} + \eta(r_t + \gamma V(S_{t+1}) - V(S_t)) \tag{1.4}$$

where η is a learning rate parameter, r_t is the reward received at timepoint t, $0 < \gamma \leq 1$ is a discounting factor to account for the fact that future rewards are less valuable than immediate rewards, and S_{t+1} is the next observable state of the environment. Temporal-difference learning models are sensitive to temporal relationships within a trial and are able to account for the timing of multiple stimuli and rewards within a trial (Niv & Montague, 2008). As in the present thesis timing within a trial was not considered to be crucial for the interpretation of the results, this extension of the reward prediction error was not incorporated into the analyses. However, early research in the animal domain, which forms the basis of the present understanding of reward-based learning, has employed this approach, as well as many fMRI studies later on (see below).

Optimizing behavior

The computational models introduced so far are aimed at describing how to optimize *predictions*. In many cases however, we are more interested in how to optimize *actions* in order to maximize the outcome, as in instrumental conditioning. Models of reinforcement learning attempt to solve this problem by either defining two modules, termed the actor and the critic, or by representing values over state-action pairs. In actor / critic models, the critic module uses temporal difference learning to estimate the value of the states of the environment. The temporal-difference prediction error that is used to train these is also conveyed to the actor module, which maintains and learns a policy, i.e. a probability distribution over all available actions at each state. In Q-learning models (Watkins, 1989) the coupling between prediction learning and action selection is more direct. In these methods, the value Q of a pair of the state S and an action a is learned (Niv, 2009). A simple form of Q-learning is implemented in the analysis of the experiment presented in Chapter 4.

Implications

The mathematical models presented in this section provide elaborate quantitative predictions about which calculations the brain might realize in order to achieve optimal decisions. They make it possible to describe putative processes by introducing hidden variables that are not readily observable, such as state values and reward prediction errors. It has to be noted that it is possible to utilize the insight gained from these models without explicitly fitting them. For example, Abler et al. (2006) kept the amount that participants could win constant and informed them about the probability of winning before each trial. In this way, they were able to calculate the value at each trial, which is assumed to be the product of these two variables (Knutson et al., 2005), as well as the reward prediction error, without estimating the free parameters of a computational model. Such an approach can be valuable as the exact estimation of free parameters for each participant can prove difficult based on the noise present in empirical data and the limited amount of trials that can be acquired within a single experiment (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006). However, this approach cannot be used to differentiate between models. The predictions of reinforcement learning models have inspired a wealth of empirical investigations providing insight to the neural mechanisms of reward-based learning. A short overview of their results is presented in the next section.

1.4.2. Neural correlates of hidden variables

Research on reward-based learning has employed a wide range of methods, from microelectrode approaches over lesion studies and pharmacological interventions to PET and fMRI. Electrophysiological recordings provide high temporal and spatial resolution, but offer information only from a restricted region. Also, due to their highly invasive nature, recordings from human participants are rarely available (exceptions are reported in Münte et al. (2007); M. X. Cohen et al. (2009)). On the other end of the spectrum, fMRI can be used to visualize activation from the whole brain, however provides a considerably lower spatiotemporal resolution and only an indirect measure of neuronal activity (see Section 1.3). The following review shortly introduces the seminal electrophysiological findings that initiated the search for neural correlates of prediction errors. It then focuses on fMRI activations observed during reward-based learning, as this is the method employed in the experiments reported in the current thesis. As estimates of the reward prediction error, the value of a stimulus and its outcome are correlated in most paradigms (Sescousse, Redouté, & Dreher, 2010; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Niv, 2009), their respective contributions are difficult to disentangle with standard fMRI methods which rely on linear regression. For the purposes of this overview therefore the activations that correlate with the different variables are reviewed in parallel.

Dopaminergic midbrain neurons

Electrophysiological recordings from the dopaminergic midbrain, i.e. the SN pars compacta and VTA, show that phasic bursts in at least a subpopulation of dopaminergic cells convey a signal corresponding closely to the reward prediction error of temporal-difference learning models (Montague, Dayan, & Sejnowski, 1996). They respond to appetitive events, such as primary rewards and reward-predicting stimuli, and do not respond when rewards are fully predicted. The omission of rewards at the predicted timepoint decreases their (low) baseline activity. This correspondence was also shown in more stringent tests of reward prediction errors, including paradigms of blocking (Waelti, Dickinson, & Schultz, 2001) and conditioned inhibition (Tobler, Dickinson, & Schultz, 2003), which further support the reward prediction error hypothesis of dopamine (for extensive reviews see Schultz, Dayan, and Montague (1997); Schultz (1998, 2000, 2006, 2007, 2010)).

Although the fMRI signal is assumed to reflect the input to a region (and processing within that region) rather than its output (see Section 1.3), several fMRI studies also report activity in midbrain at or near the SN / VTA in response to reward prediction and the prediction error (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; D'Ardenne et al., 2008; Dreher, Kohn, & Berman, 2006; O'Doherty, Deichmann, Critchley, & Dolan, 2002; O'Doherty, Buchanan, Seymour, & Dolan, 2006; Schott et al., 2007; Wittmann et al., 2005), with however no substantial decreases in activation when expected rewards fail to occur (Haber & Knutson, 2009).

Striatum

More consistently than midbrain activations, activations in dopaminergic target areas, including the striatum, have been reported. Activation in the ventral striatum, i.e. the ventral putamen and nucleus accumbens, correlates with the prediction error during Pavlovian conditioning, while activations in both the dorsal and ventral striatum correlate with the prediction error during instrumental conditioning (O'Doherty, 2004). On this basis, a dorsal / ventral dissociation of the striatum has been suggested analogously to the actor / critic distinction in reinforcement learning. Indeed, some fMRI studies of action-contingent reward learning report dorsal striatal activation (Balleine, Delgado, & Hikosaka, 2007; Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Delgado, Miller, Inati, & Phelps, 2005; Elliott, Newman, Longe, & William Deakin, 2004; Haruno et al., 2004; Haruno & Kawato, 2006; Knutson, Fong, Adams, Varner, & Hommer, 2001; Tricomi, Delgado, & Fiez, 2004).

The area most consistently observed to be activated in response to reward anticipation and the prediction error is the ventral striatum (e.g. Abler et al. (2006); Bjork and Hommer (2007); Gläscher,

Daw, Dayan, and O'Doherty (2010); Lin, Adolphs, and Rangel (2012); O'Doherty, Dayan, Friston, Critchley, and Dolan (2003); Pagnoni, Zink, Montague, and Berns (2002); Preuschoff, Bossaerts, and Quartz (2006); Knutson et al. (2001); Niv, Edlund, Dayan, and O'Doherty (2012); Rodriguez et al. (2006); Schönberg, Daw, Joel, and O'Doherty (2007); Yacubian et al. (2006)). FMRI activation correlating with parameters predicted by reinforcement learning models has been observed for a wide range of rewards, ranging from juice to smiles / positive social cues (Lin et al., 2012), but also for punishments like cutaneous electrical stimulation (Menon et al., 2007). While some studies report positive and negative reward prediction errors to be associated with increased and decreased activity, respectively (see e.g. Abler et al. (2006); Pessiglione et al. (2006); Schönberg et al. (2007)), others have observed increases in striatal fMRI activations in response to errors in the prediction of aversive events (see e.g. Jensen et al. (2007); Menon et al. (2007); Seymour et al. (2004)). The source of these positive activations in response to negative prediction errors is unclear, but valence-independent dopaminergic processes (Metereau & Dreher, 2012), non-dopaminergic opponent processes signaling aversive prediction errors (Seymour, Daw, Dayan, Singer, & Dolan, 2007), and the potential inability of fMRI to distinguish between excitatory and inhibitory input (Niv, 2009) have been proposed.

The amygdala and lateral habenula are also among the subcortical structures that have been implicated in reinforcement learning, however mainly in the context of aversive learning. As this is not a focus of the present thesis, the reader is referred to Haber and Knutson (2009) for a recent review of their potential function.

Prefrontal cortex

Reward signals are observable not only in subcortical structures but also in most parts of the cortex (Vickery, Chun, & Lee, 2011). Three prefrontal areas seem to play a special role in reinforcement learning: the ventromedial prefrontal cortex (vmPFC), orbitofrontal cortex (OFC), and dorsomedial prefrontal cortex (dmPFC) / dorsal anterior cingulate cortex (dACC). They project directly to the ventral striatum and are connected with it in the reward circuit (Haber & Knutson, 2009). Functional MRI activation in the vmPFC and adjacent medial OFC has been shown to correlate with an abstract representation of value (Haber & Knutson, 2009). Some controversy exists about the functional division between medial and lateral OFC regions, however there are indications that the vmPFC, and possibly the adjacent medial OFC, process diverse and abstract rewards, while the lateral OFC is more specialized in its responses to different types of reinforcers (Sescousse et al., 2010) (for reviews see Haber and Knutson (2009); Kringelbach and Rolls (2004); O'Doherty (2007); Rushworth, Noonan, Boorman, Walton, and Behrens (2011); Wallis (2007)). Lastly, the dmPFC / dACC is involved in reward-guided learning with an emphasis on action selection. Due to the location of the dmPFC, its activation can also be assessed using non-invasive electrophysiological measurements from the scalp (Debener et al., 2005), and there is evidence that two electroencephalography (EEG) components associated with it, the error-related negativity (ERN) and feedback-related negativity (FRN), reflect several characteristics of the prediction error signal (Holroyd & Coles, 2002; Holroyd, Yeung, Coles, & Cohen, 2005; M. X. Cohen & Ranganath, 2007). The dmPFC / dACC, and especially the rostral cingulate zone (RCZ), participate in monitoring response conflicts and uncertainty, performance outcomes and response errors (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). One important dissociation of activations in response to errors and rewards can be observed between the dmPFC and the ventral striatum: the dmPFC is activated by all errors regardless of their effect on outcome, while the ventral striatum is activated by reward regardless of the correctness of the preceding action (de Brujin, de Lange, von Cramon, & Ullsperger, 2009). Additionally, dmPFC / dACC activation indicates when action values should be adjusted, and predicts switches in action selection (Rushworth & Behrens, 2008), emphasizing the central role of these areas for guiding future actions based on their past outcomes.

Overview

The main areas in which fMRI activation has been shown to correlate with variables from reinforcement learning models are the VTA / SN, the basal ganglia (mainly the striatum), subcortical limbic structures including the amygdala, and the PFC (mainly the medial prefrontal cortex (mPFC) and the OFC). In a more vigorous approach than can be achieved by examining correlation coefficients between model-derived hidden variables and neural activation data, Rutledge, Dean, Caplin, and Glimcher (2010) tested an axiomatic model in eleven brain areas to assess whether the recorded activation during a reward learning task possesses the necessary and sufficient properties of any reward prediction error signal. They were able to provide evidence for this in the nucleus accumbens, caudate, putamen, amygdala, mPFC, and posterior cingulate cortex.

1.4.3. Concluding remarks

The reward prediction error hypothesis of dopamine is not unchallenged (see e.g. (Berridge, 2007; Redgrave, Gurney, & Reynolds, 2008; Vitay, Fix, Beuth, Schroll, & Hamker, 2009)) and, without further extensions, the models presented here are not able to account for more complex phenomena observed in learning, like dopaminergic responses to novelty (Wittmann, Bunzeck, Dolan, & Düzel, 2007), the sensitivity to risk (Dayan & Niv, 2008; Niv et al., 2012), or the ability to learn from rewards that were foregone (Lohrenz, McCabe, Camerer, & Montague, 2007). However, even in their most simplistic form reinforcement learning models provide a parsimonious account of many behavioral phenomena, as well as of the pattern of neural activations observed using fMRI. Currently, the prediction error hypothesis has inspired the most extensive body of experimental research on reward-based learning, and models derived from it have proven to be able to predict neural activations across species and for a wide range of reinforcing stimuli and experimental paradigms. Therefore, the studies presented in this thesis rely on this framework to both interpret the observed findings and to compare the fMRI activations observed during learning in the absence of explicit reinforcement with those observed during reward-based learning.

1.5. Open questions addressed by this thesis and outline of the experiments

As summarized in Sections 1.2-1.4 within the last decade considerable advances have been made in understanding the neural bases of human reward-based learning and category learning. However, as these two lines of research have mainly developed in parallel, some open questions remain especially in the fields where they overlap. The experiments presented in this thesis are designed to address the central question of the influence of the nature of outcome information on the neural substrates of category learning.

As noted in Section 1.2, feedback plays a central role in complex categorization tasks where information has to be integrated across multiple experiences (Shohamy et al., 2008) and / or within the same stimulus (Ashby & Maddox, 2005). This has been postulated to be due to the fact that learning in subcortical structures, which support complex category learning, is dependent on dopamine-mediated reward signals from dopaminergic midbrain neurons (Seger et al., 2010; Ashby & Maddox, 2005; Seger, 2008). Here, the implicit assumption is made that cognitive feedback, i.e. information about the correctness of an answer, activates the reward system in a similar way as explicit rewards. Most studies on learning in humans were conducted using fMRI, which is not able to index dopamine release directly. However, based on data from pharmacological, neuropsychological and PET approaches (see Section 1.3), currently it has to be assumed that in reward-based paradigms there is a relationship between the fMRI signal, dopaminergic activity in the midbrain, and dopamine release in the ventral striatum. Independent of the actual source of the fMRI signal, vast evidence exists that structures of the dopaminergic system,

especially the ventral striatum, show fMRI activation in response to a wide range of reinforcing stimuli, including cognitive feedback (Aron et al., 2004; Rodriguez et al., 2006), in a pattern consistent with reinforcement learning theory (see Section 1.4).

However, similar activation patterns in different tasks are not a sufficient proof that the same processes are implicated (Poldrack & Willingham, 2006), and can only generate hypotheses about potential differences in processing. In the first experiment presented in this thesis (Chapter 3) we therefore directly compared the effects of explicit reward, operationalized as monetary gain, to the effects of cognitive feedback, operationalized as information about the correctness of an answer. To this end, two parallel versions of an information-integration category learning task were developed. As internal error signals are also known to activate structures of the reward circuit (Holroyd et al., 2004), in the second experiment (Chapter 4) this approach was expanded by using a version of the same paradigm in which no external feedback was provided. The participants were questioned about their internal signals on correctness and the neural correlates of these internal signals were examined in terms of reinforcement learning theory. In a final step, the lasting effects of monetary reinforcement and cognitive feedback on the categorical representation of stimuli after training were compared (Chapter 5). All experiments were performed on young healthy human participants using fMRI. Specific hypotheses for each experiment are detailed in their respective introduction. In general, we assumed that category learning in the absence of explicit rewards is supported by the dopaminergic system in a similar way as it is in reward-based learning. However, quantitative and qualitative differences in involvement were expected, which are specified in the context of the presented experiments.

2 General Methods

Given that there is a considerable overlap of methodology in the three experiments presented in Chapters 3-5, this chapter summarizes all common methods to avoid repetitions. Methodological aspects which were specific to each experiment are detailed in their respective methods section.

2.1. Participants

For all experiments young healthy adults of both genders were recruited as participants from the Otto-von-Guericke University community. None of them reported a history of drug abuse, psychiatric diseases or neurological injuries. All participants were without pathological findings on a psychiatric screening questionnaire (SCL-90-R; Franke (1995)), right-handed as confirmed by the Edinburgh Handedness Inventory (Oldfield, 1971) and reported normal or corrected to normal vision. Written informed consent was obtained prior to the experiments in accordance with the protocols approved by the ethics committee of the Otto-von-Guericke University.

2.2. Behavioral task

All categorization tasks utilized in the present work were developed based on the randomization technique of general recognition theory (Ashby & Gott, 1988). As within this family of categorization tasks the categories are defined by bivariate normal distributions within an arbitrary two-dimensional stimulus space, extensive control over all important aspects of the categories is provided. The technique allows adjusting the category structure, the maximum possible accuracy rates, and the shape of the categorization rule to fit the current experiment. In the presented studies three types of stimuli were used: circles with an opening of 30 ° width, two parallel lines, and circular square wave gratings. The dimensions defining the stimulus space were linewidth and orientation for the circle and line stimuli, and spatial frequency and orientation for the gratings. By choosing bivariate normal distributions with a non-zero covariance, all category boundaries were constructed to be oblique to the dimensions spanning the stimulus space. This requires participants to integrate information from both stimulus dimensions predecisionally, which is known to lead to a gradual acquisition of category knowledge and to be more dependent on training than the acquisition of easily verbalizable rules (Ashby & Maddox, 2005).

For each trial a category was chosen randomly with equal baserates, and an ordered pair of values was sampled randomly from the appropriate bivariate normal distribution. Within each experimental block the sample mean-vector and variance-covariance matrix of the sampled stimuli were transformed to equal the predefined distribution specifying the category (exact parameter values are listed in Tables 3.1, 4.1, and 5.1). To construct the stimuli, the values obtained in this way were transformed to the dimensions specifying the stimulus space (orientation and linewidth or spatial frequency) using

transformation parameters that were chosen in an attempt to equalize the salience of both dimensions. The resulting category structures along with sample stimuli for each category and experiment are presented in Figure 3.1, Figure 4.1 and Figure 5.1.

2.3. Model-based analysis of the behavioral categorization data

As Experiments 1 and 2 were conducted to examine the neural correlates of the category learning process, fMRI data was acquired during learning, i.e. when the participants' performance was still suboptimal. Previous studies have shown that tasks which are solved by applying a simple verbal rule recruit different neural structures as information-integration tasks (Ashby & Maddox, 2005). Therefore, it is crucial for the interpretation of the results to determine which strategy each participant used to solve it. Since this information cannot be provided by measures of overall performance, three different types of models were fit to the behavioral data based on the location of the responses in the two-dimensional stimulus space using procedures established by Maddox and Ashby (1993).

2.3.1. Unidimensional models

Unidimensional models assume that the participant sets a decision criterion on a single stimulus dimension (i.e. either the linewidth or orientation of the stimuli) and then makes an explicit decision about the value of the stimulus on the relevant dimension. An example for a unidimensional rule is: "Respond A if the circle is open to the left, respond B if it is open to the right". These models have two free parameters: the decision criterion on the relevant dimension and the variance of internal noise around the decision bound. In the tasks employed in Experiment 1 and 2 a participant using the best unidimensional rule will respond correctly in 67% of the trials.

2.3.2. Conjunction models

Conjunction models assume that the participant uses a criterion on each stimulus dimension and decides about the value of the stimulus on both dimensions. Subsequently, a conjunctive rule based on the outcome of these decisions is applied. Two conjunctive models were fit to each task. The two rules fit to the data of the task with circle stimuli and a positive slope of the decision bound are presented as examples of conjunction rules:

(1) "Respond A if the circles are open to the right and are thick, otherwise respond B."

(2) "Respond B if the circles are open to the left and are thin, otherwise respond A."

These models have three free parameters: a decision criterion on each dimension and the variance of internal noise. In the tasks presented in Experiment 1 and 2 a participant using the optimal conjunctive rule will respond correctly in 80% of the trials.

2.3.3. Information-Integration models

Information-integration models assume that the participant uses a decision criterion that integrates information from both dimensions predecisionally. In the present work all optimal decision bounds were linear. Within the information-integration model the slope of the decision bound was fixed to +/-1, leaving two free parameters (intercept and noise variance). Given the non-overlapping category distributions employed in this thesis, the use of an optimal information-integration rule will result in 100% correct answers.

2.3.4. Model estimation and selection

Each of the models described above was fit separately to each dataset. The free parameters were estimated using the method of maximum likelihood (Ashby, 1992) and the Bayesian information criterion (BIC) (Schwarz, 1978) was used as a goodness of fit statistic for model selection:

$$BIC = rlnN - 2lnL$$

where r is the number of free parameters, N is the sample size and L is the likelihood of the model given the data. The BIC penalizes a model for additional free parameters so that a set of models can be compared directly.

2.4. FMRI Image acquisition

Functional magnetic resonance imaging data was acquired on a Siemens MAGNETOM Trio (Erlangen, Germany) whole body 3T magnetic resonance imaging (MRI) scanner equipped with an 8 channel head coil. First, structural images of the brain were recorded using a T1-weighted magnetization-prepared rapid acquisition gradient echo (MP-RAGE) sequence with a field of view (FoV) of 256 mm and 192 sagittal slices with 1 mm^3 isotropic voxels (time to repeat (TR) = 2500 ms, time to echo (TE)= 4.77 ms, inversion time (TI) = 1100 ms, flip angle (FA) = 7°, slice gap = 0.5 mm). Subsequently, functional volumes were obtained using a whole-brain $T2^*$-weighted echo planar imaging (EPI) sequence. The parameters for the functional measurements were: TE = 30 ms, TR = 2000 ms, slice thickness = 3 mm, slice gap = 0.3 mm, number of slices = 32 (interleaved order), FA = 80°, FoV = 192 mm × 192 mm and a matrix size of 64 × 64, resulting in an inplane resolution of 3 mm × 3 mm.

3 Experiment 1: Comparing the neural basis of monetary reward and cognitive feedback during information-integration category learning

The results of this experiment were first published in: Daniel, R., & Pollmann, S. (2010). Comparing the neural basis of monetary reward and cognitive feedback during information-integration category learning. *The Journal of Neuroscience*, *30*(1), 47–55.

3.1. Introduction

Learning which action leads to the most beneficial outcome in a given situation is one of the central components of adaptive behavior. The dopaminergic system with its projections to striatal and medial prefrontal areas is known to play a crucial role in reward learning (O'Doherty, 2004; Schultz, 2006). In humans, it is often studied by using gambling paradigms, where participants learn probabilistic stimulus-reward contingencies by trial and error, the reward being earnings in money (see e.g. Abler et al. (2006), Dreher et al. (2006) or B. W. Smith et al. (2009)). However, there are indications that the dopaminergic system is also involved in tasks where only cognitive feedback is provided (Aron et al., 2004; Rodriguez et al., 2006).

Behavioral studies of category learning with cognitive feedback suggest that tasks which force participants to rely on gradually acquired stimulus-outcome contingencies are sensitive to the nature and timing of feedback (Maddox et al., 2003, 2008). During trial and error learning in these tasks striatal areas are activated (Cincotta & Seger, 2007; Poldrack et al., 2001) and patients with Parkinson's disease, which is characterized by the loss of dopaminergic input to the striatum, are impaired (Filoteo, Maddox, & Davis, 2001). Results therefore indicate that the functional neuronal substrates underlying learning based on reward and cognitive feedback are very similar in certain task domains.

However, differences have to be expected as the dopaminergic system is known to respond differentially to rewards of different magnitude and value in the insula, amygdala, orbitofrontal cortex and striatum (Gottfried, O'Doherty, & Dolan, 2003; B. W. Smith et al., 2009; Tobler, O'Doherty, Dolan, & Schultz, 2007). Also, subcomponents of the striatum coding motivational aspects, like the nucleus accumbens, are assumed to respond to the modulation of reward characteristics, while subcomponents involved in executive processes, like the caudate head, should be less affected. Additionally, it is known that providing external rewards can undermine intrinsic motivation, i.e. motivation that is based on interest in and enjoyment of a task. Motivational states have been shown to modulate the fMRI response in the striatum (Mizuno et al., 2008; Murayama, Matsumoto, Izuma, & Matsumoto, 2010), with less

activation in response to feedback if intrinsic motivation has been undermined by providing monetary rewards. Therefore the striatal activation in response to the expectation and receipt of cognitive feedback and monetary reward might not only differ quantitatively, i.e. in the height of the response, but also qualitative differences which are mediated by motivational states have to be expected.

To test these assumptions we conducted an fMRI experiment comparing the effects of cognitive feedback and reward in an information-integration category learning task (Ashby et al., 1998). Each participant performed two parallel versions, in one of which correct answers were rewarded with a monetary gain, while in the other only information about the correctness of the answer was provided. Although there are indications that both negative and positive feedback contribute to information-integration learning (Ashby & O'Brien, 2007), the dopaminergic substrates of reward learning are well established, while research on avoidance learning mainly focused on the amygdaloid-hippocampal basis of fear conditioning (LeDoux, 2003). Therefore in the present study no monetary punishment was delivered. As with training the dopaminergic response is known to shift backwards in time from the reward to the reward predicting stimulus (Schultz, 2000), we expected to find an effect of the reward manipulation already during stimulus presentation. If reward manipulation also has an influence on the prediction error signal (Schultz, 2007), differential responses between the tasks to both negative and positive feedback are predicted, as long as performance is not perfect. We assessed the motivational state of each participant using self report measures (Ryan, 1982) and anticipated that the effect of monetary reward is primarily predicted by measures of extrinsic motivation, while the effect of cognitive feedback is more responsive to measures of intrinsic motivation.

3.2. Methods

3.2.1. Participants

Sixteen participants with an average age of 23.1 years [range = 18-29; SD = 2.9; 10 females] participated in the experiment. They received an average payment of € 29.3 [range = 26-33; SD = 2.5] based on their performance.

3.2.2. Stimuli

Two sets of stimuli were presented in the present experiment, either circles with an opening of 30° width or two parallel lines, both in white on black background. All stimuli varied on two dimensions, linewidth and orientation, and categories were specified based on the location of the stimuli within this two-dimensional perceptual space (see Figure 3.1). Two parallel category structures for each set of stimuli were specified by means of rotating the category boundary by 90°, resulting in the same within category scatter, difference between category means and absolute value of covariances of the two dimensions in both category structures. The parameters of the bivariate normal distributions specifying the four categories are summarized in Table 3.1. Values for x and y were sampled randomly from these distributions and used to determine linewidth ($x/2$ pixels) and orientation ($y \times \pi/1200$ radians) of the stimuli. As linewidth only increased to the center of the stimuli, the overall dimensions of the images were not changed by the manipulation. The optimal decision bound for both category structures is depicted along with examples for both stimulus sets in Figure 3.1.

3.2.3. Procedure

Each participant performed two tasks: In the first task, one type of stimuli (lines or circles) was presented with one category structure (positive or negative slope of the optimal decision bound) and in the second task, the other type of stimuli was presented with the other category structure. One of these two tasks was rewarded with € 0.20 for each percent of right answers, resulting in eight possible task combinations (2 (stimulus type) x 2 (category structure) x 2 (first or second task rewarded)), each of which was

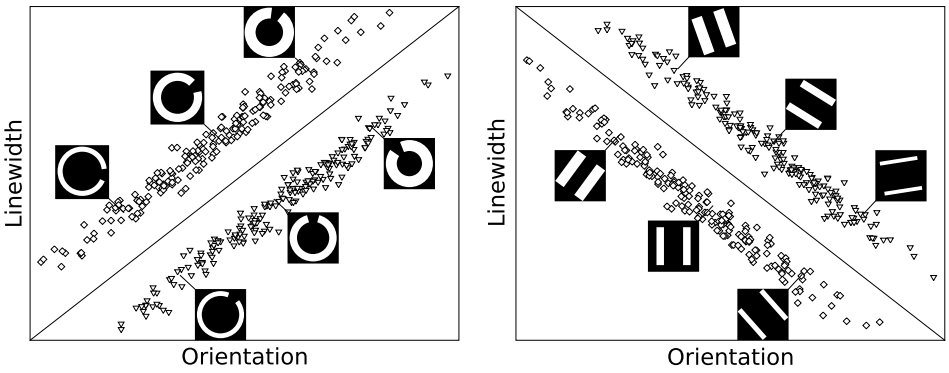

Figure 3.1. *Category structures and sample stimuli.* Each square denotes the orientation and linewidth of a stimulus from category A, each triangle those of a stimulus from Category B. The lines represent the optimal decision bound. Two types of category structure were presented, one with a positive slope of the optimal decision bound and one with a negative slope. For both types of stimuli used in the experiment, circles and lines, examples of three stimuli from each category are shown. Both types of stimuli were used with both types of decision bounds. Note that 0° does not correspond to a horizontal alignment of the stimuli to make the verbalization of a categorization rule more difficult.

Table 3.1. *Category distribution parameters*

	Category A					*Category B*				
Slope[a]	μ_x	μ_y	σ_x	σ_y	cov_{xy}	μ_x	μ_y	σ_x	σ_y	cov_{xy}
Positive	500	700	44100	44100	43500	700	500	4410	44100	43500
Negative	500	500	44100	44100	-43500	700	700	4410	44100	-43500

μ = mean for each dimension; σ = variance for each dimension; cov = covariance between dimensions.
[a] of the optimal decision bound

presented to two of the 16 participants. Participants were instructed to learn about the two categories in each task by relying on the feedback they would receive after each decision and were informed that perfect performance was possible. Since we were not interested in the early performance on this task, while participants might still employ suboptimal verbal strategies, all participants were trained to criterion (80% correct answers within a single block of 50 trials) on the day before the fMRI session. The training was performed in a dimly lit room using Presentation (Neurobehavioral Systems Inc., Albany, CA). All participants were presented with the two tasks in alternating blocks of 50 trials with equal base rates for both categories. Each trial consisted of the presentation of a stimulus spanning a visual angle of 12° at the center of the screen for 2 s. Participants were requested to make a decision about category membership during stimulus presentation. After a random delay that was sampled from an exponential distribution with a mean of 2 s [range 0.5 - 6 s] they received both auditory and visual feedback. Auditory feedback was provided via a tone of 0.25 s duration and had a frequency of 900 Hz for right and of 350 Hz for wrong answers, while visual feedback was presented for 1 s at a visual angle of 10°. Positive visual feedback consisted of a filled green circle in the cognitive feedback task and of a picture of a 20 cent coin in the monetary reward task. Negative feedback was indicated in both tasks by a filled red circle. When the participant failed to respond, a filled yellow circle was presented. Except the intertrial interval (ITI), which was sampled from an exponential distribution with a mean of 3.5 s [range 0.5 - 8 s], all parameters were equivalent to the fMRI session (see Figure 3.2). Training ended independently for each task after the criterion of 80% correct answers in a single block was reached. Two participants did not reach the criterion after 5 blocks and were therefore not included in the study. Training was followed by the fMRI

experiment on the next day. During the fMRI experiment participants completed the two tasks they had trained alternatingly in four blocks. Each block incorporated 50 trials and lasted approximately 11 min. All stimuli originated from the same bivariate distributions as the training stimuli, but were sampled independently of them. To exclude the possibility that activation differences between the rewarded and unrewarded task are exclusively due to differing visual stimulation, for half of the participants the 20 cent coin signaling a correct answer in the rewarded task was replaced by a green circle with the instruction that they would nevertheless gain € 0.20 for each correct answer. The single trials were separated by a variable ITI sampled from an exponential distribution with a mean of 6 s [range 1-12 s]. An overview of the experimental procedure is given in Figure 3.3.

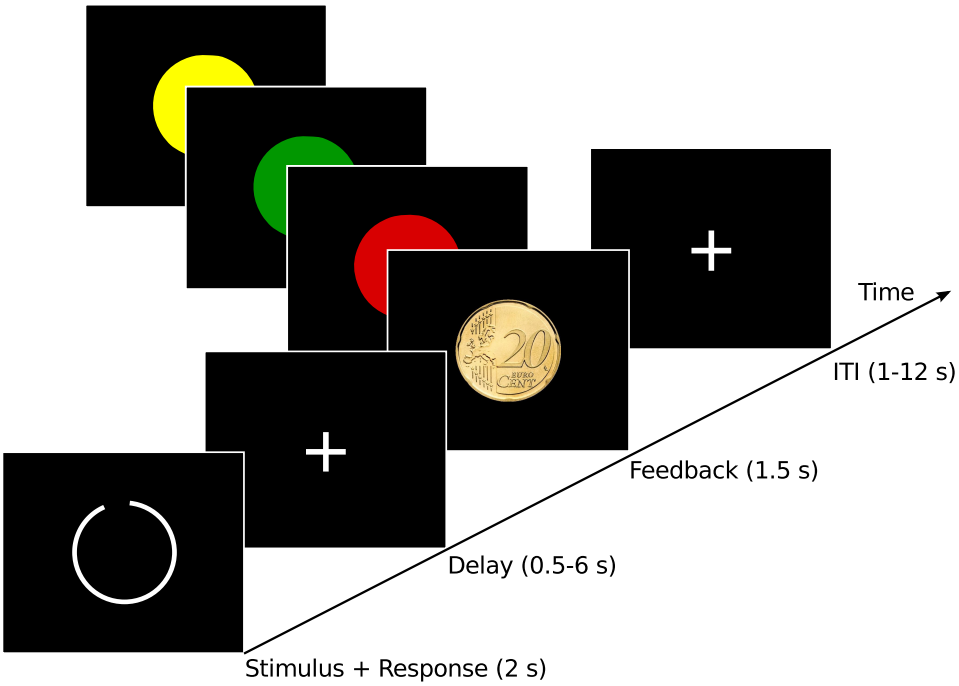

Figure 3.2. *Trial structure.* Each trial started with the presentation of a stimulus for two seconds. Participants were instructed to respond during this period by pressing of one of two buttons. The stimulus was followed by a delay that was randomly sampled from an exponential distribution with a mean of 2 s (range 0.5-6 s), after which feedback was presented for 1.5 s. Positive feedback was given by showing a green circle, or, in the rewarded condition, by a 20 cent coin, and a high tone. Negative feedback consisted of a red circle and a low tone. If the participant failed to respond a yellow circle was presented together with the low tone. Trials were separated by an interval that was randomly sampled from an exponential distribution with a mean of 6 s (range 1-12 s).

After completing the testing session all participants filled out a questionnaire based on the post-experimental scale of the Intrinsic Motivation Inventory (IMI) (Ryan, 1982; McAuley, 1989) for both tasks. The included subscales were *Interest / Enjoyment*, *Perceived competence*, *Perceived choice*, *Pressure / Tension*, *Effort / Importance*, and *Value / Usefulness*.

3.2.4. Pilot testing of the task versions

To ensure that the four task versions (2 sets of stimuli × 2 category structures) do not differ in terms of learning speed and error rates, a pilot study with 17 participants was conducted. Participants were recruited from the same population as the participants the fMRI study, and the procedures were equivalent to those of detailed above with the sole exception that both sessions were conducted in

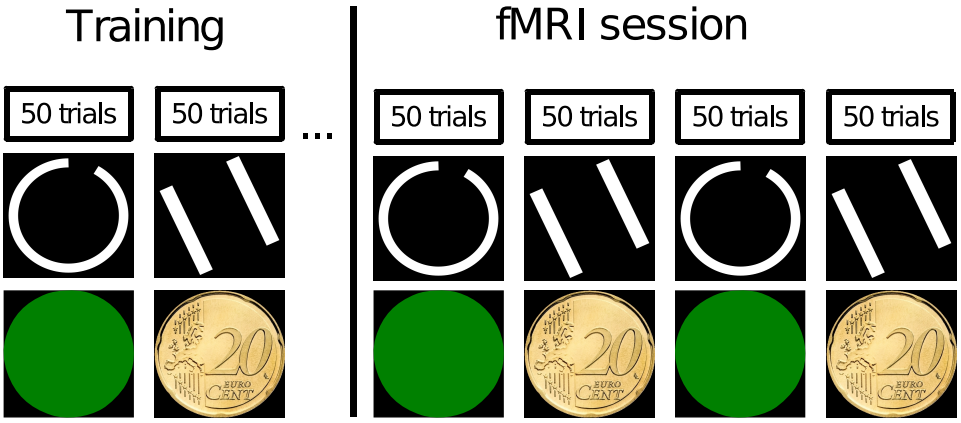

Figure 3.3. *Session structure.* Each participant was trained on both tasks on the day before the fMRI session. Whether the first task was rewarded or not, whether it contained circle or line stimuli, and whether the optimal decision bound had a positive or negative slope, was randomized across participants. Training ended independently for both tasks after the participant reached an accuracy rate of 80% within a single block. During the fMRI session, the two trained tasks were presented alternatingly in four blocks of 50 trials each.

the same room. Detailed results are provided in Appendix A.1; importantly, no significant differences between task versions were observed at an alpha level of .2.

3.2.5. FMRI image acquisition and image preprocessing

The parameters for image acquisition are detailed in Section 2.4. An overall of 1360 $T2^*$-weighted EPI volumes were obtained in a single run. The functional images were preprocessed with the statistical parametric mapping software SPM5 (Wellcome Department of Cognitive Neurology, London, UK). Preprocessing included slice timing correction using the first slice as reference and three-dimensional motion correction, i.e. rigid body realignment to the mean of all images. The six estimated movement parameters were saved and later included in the statistical analysis. It was ensured that head movement was below 3 mm and 3° for each participant. Images were normalized to Montreal Neurological Institute (MNI) space (Evans et al., 1993) using the standard EPI template of SPM5. The data was spatially smoothed using a Gaussian filter of 6 mm full width at half maximum (FWHM) and a temporal highpass filter of 1/128 Hz was applied to remove low frequency confounds.

3.2.6. Statistical analysis

The regressors for within-participant modeling were convolved with a canonical model of the hemodynamic response function (HRF) as implemented in SPM5 and represented right and wrong answers of each participant for both the period of *expectation* (i.e. stimulus presentation and response) and *feedback* presentation. An additional regressor for the *delay* between *expectation* and *feedback* was included and all regressors were fit to the data separately for the rewarded and not rewarded condition using the general linear model (GLM). Contrast images of the condition-specific estimates for each participant were then submitted to the second level group analyses with participant as the random-effect variable. A $2 \times 2 \times 2$ analysis of variance (ANOVA) was conducted with the factors *trial part* (stimulus / feedback), *condition* (rewarded / non-rewarded) and *success* (right / wrong answer). All fMRI activation maps were thresholded at $p < .05$ (spatial extent > 5 contiguous voxels) and all reported clusters survived correction for multiple comparisons at the whole brain level using family-wise errorrate (FWE) correction (Worsley, Marrett, & Neelin, 1996). As an analysis comparing the group of participants receiving visual positive feedback in the rewarded condition in form of a coin with those receiving it in form of a green circle did not reveal any significant differences, the data of both groups were collapsed.

3.3. Results

3.3.1. Behavioral results

Accuracy measures

Participants fulfilled the criterion of 80% correct answers after an average of 131.25 [$SE = 20.09$] training trials in the rewarded task and after an average of 129.69 [$SE = 17.56$] training trials in the unrewarded task. An ANOVA of the behavioral data collected during the fMRI session with repeated measures on *condition* (rewarded / not rewarded) and *block* (first / second block of the experiment) showed a significant main effect of *block* [$F(1, 15) = 5.5, p < .05$] and a significant interaction of *condition* × *block* [$F(1, 15) = 5.5, p < .05$]. Post-hoc paired samples T-tests revealed that the interaction was due to the fact that for the rewarded condition error rates significantly decreased from a higher base level from the first to the second block [$T(15) = 3.3, p < .05$], while they did not change in the unrewarded condition [$T(15) = 0.5, n.s.$]. The average error rates were 23.7% (SE = 1.6) in the monetary reward condition and 23.4% (SE = 1.4) in the cognitive feedback condition.

Model-based analysis

An analysis of the individual categorization strategy was calculated based on the location of each participant's responses in the two dimensional stimulus space as described in Section 2.3. The models were separately fit to each of the 32 datasets (two tasks for each of the 16 participants). Results suggest that for nine of the 16 participants at least one data set is best fit by the information-integration model which assumes that the participant applied the optimal decision bound depicted in Figure 3.1 and therefore integrated information from both dimensions. Given the few data points the model fits are based on, these results have to be interpreted with care. Nevertheless, the modeling results indicate that not all participants engaged the procedural-learning based system mediated by subcortical structures. The influence of individual strategies on brain activation is addressed in Section 3.3.2. More detailed results on model fit are provided in Appendix A.2.

Questionnaire data

ANOVAs were performed with repeated measures on condition (rewarded / not rewarded) with the subscales of the post-experimental motivation inventory as dependent measures. No significant effects were observed.

3.3.2. Functional imaging results

Effect of monetary reward versus cognitive feedback

Comparing the effect of monetary reward and cognitive feedback revealed a significantly higher activation for monetary reward bilaterally in the nucleus accumbens (MNI: x = 9, y = 6, z = -9; Max. T: 6.59 and x = -6, y = 0, z = -6; Max. T: 5.52) during *expectation* (see Figure 3.4). No areas were significantly more activated during *expectation* in the cognitive feedback task nor were any differences observed during the receipt of feedback in both tasks. When the effects of *feedback* in the rewarded and the cognitive feedback task were compared separately for positive and negative feedback, no significant differences were observed.

Examining the nucleus accumbens activation

Correlation of the BOLD signal with questionnaire data To assess the effect of the participants' motivational state on activations, the estimated signal change (percent) within the nucleus accumbens was extracted to both monetary reward and cognitive feedback during *expectation* using MarsBar

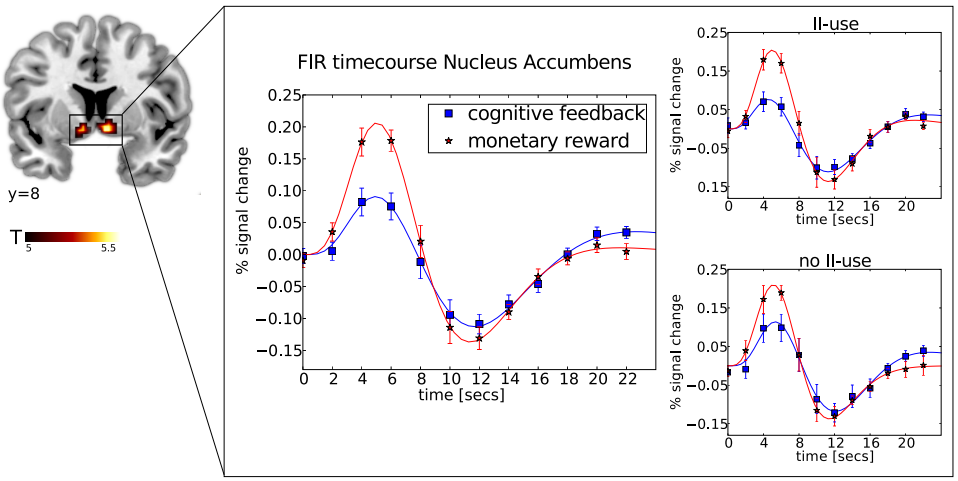

Figure 3.4. *Effect of monetary reward.* Activation in the contrast of monetary reward minus cognitive feedback during stimulus presentation. The timecourse represents the finite impulse response to both monetary reward and cognitive feedback during stimulus presentation, extracted using MarsBar and an anatomical ROI of the nucleus accumbens from the Harvard-Oxford subcortical structural atlas. For each participant individual functional ROIs within this anatomical ROI were defined based on the areas where the main effect of stimulus presentation exceeded an uncorrected threshold of $p < 0.1$. Error bars represent the standard error of means. Results of this analysis are also plotted separately for participants with use of the optimal decision bound in at least one condition (II use) as compared to those with no information-integration use (no II use). A significant peak activation difference between the task conditions is only observed in the group of participants with II use. The map was thresholded at $p_{FWE} < .05$; the left hemisphere is presented at the left.

(marsbar.sourceforge.net) and an anatomical region of interest (ROI) of the nucleus accumbens from the Harvard-Oxford subcortical structural atlas as implemented in the FMRIB Software library (FSL; www.fmrib.ox.ac.uk). For each participant individual functional ROIs within this anatomical ROI were defined based on the areas where the main effect of stimulus presentation exceeded an uncorrected threshold of $p < .1$. This data was then submitted to a multiple stepwise regression with the individual scores on the six motivation questionnaire subscales. In each task a single predictor was sufficient for a significant model: in the cognitive feedback task *Perceived competence* predicted activation within the nucleus accumbens, [$\beta = .023$, $p < .05$], while in the monetary reward task it was predicted by *Pressure / Tension* [$\beta = .022$, $p < .05$].

Influence of individual strategies As the model based analyses indicate that not all participants employed the optimal information-integration decision bound, the data set was split into participants whose data was fit best by an information-integration model in at least one task and those participants for whom this was not the case. A stronger BOLD signal in response to the expectation of monetary reward as compared to cognitive feedback in the nucleus accumbens was only present in the group of participants employing an information-integration rule [$T = 2.93$, $df = 8$, $p < .05$]. For participants putatively employing a rule-based strategy (i.e. a conjunctive rule) peak activations in the nucleus accumbens did not differ between the expectation of monetary reward and cognitive feedback [$T = 1.67$, $df = 6$, $p < .15$]. (see Figure 3.4).

Successful and unsuccessful categorization versus fixation

To ensure that the lack of differential activation in response to cognitive feedback and monetary reward in other areas than the nucleus accumbens was not due to a general failure of our paradigm to activate dopaminergic structures, both successful and unsuccessful categorization were compared separately to fixation. We observed widespread common bilateral cortical activations extending from the visual

cortices (Brodmann area (BA) 17/18/19) ventrally to the posterior temporal cortex (BA 37) as well as dorsally to the posterior parietal cortex (BA 39/40/7). Also superior frontal (BA 6), ventrolateral prefrontal (BA 45), medial prefrontal (BA 24/32) and anterior insular activations were observed in both contrasts. Signal decreases were found for both contrasts in medial orbitofrontal areas (BA 9/10/11), in the angular gyrus (BA 39), and in the middle temporal gyrus (BA 21).

Activations during successful categorization were observed in the parahippocampi, the thalamus, the head of the caudate and the anterior pallidum. Moreover, a cluster of activation was found bilaterally in the midbrain (MNI: x = 9/-9, y = -15, z = -15) at or near the substantia nigra, both for successful and unsuccessful categorization. Further activations detected when comparing unsuccessful categorization with baseline were located bilaterally within the parahippocampi, the thalamus and a cluster in the left anterior pallidum, but not the caudate (see Figure 3.5 A and B).

Successful versus unsuccessful categorization

When comparing successful and unsuccessful categorization with fixation, we observed different activation patterns in our areas of interest, i.e. subcortical dopaminergic projection sites. We therefore directly compared the two. Bilateral activations within the putamen were observed along with a cluster in the left posterior parietal cortex (BA 40) to show significantly higher activations during successful as compared to unsuccessful categorization. No areas showed significantly higher activations for unsuccessful as compared to successful categorization (see Table 3.2 and Figure 3.5 C).

Positive and negative feedback versus fixation

As even after training performance on the task was not perfect, both positive and negative feedback are not fully predicted and therefore are expected to elicit a prediction error signal. To examine this effect, we compared both events separately to fixation. Positive feedback significantly activated cortical areas in the (pre-)cuneus, middle temporal gyrus (BA 20/21), angular gyrus (BA 39), posterior parietal cortex (BA 7/40) and superior frontal areas (BA 9/10). Deactivations were detected bilaterally in the anterior insula. Activations in this contrast included bilaterally all parts of the caudate (head, body and tail) as well as both hippocampi. When comparing negative feedback with fixation, no subcortical activations were found at the chosen threshold. Cortical activations included left prefrontal cortex extending from dorsolateral to frontopolar areas (BA 9/10/45/46), the precuneus, bilaterally the middle/superior temporal gyri (BA 21/22) as well as a cluster in the right the posterior parietal cortex (BA 40). Also both anterior insulae, the right supplementary motor area, and the dorsal anterior cingulate cortex (BA 32) were activated (see Figure 3.5 D and E).

Positive versus negative feedback

The direct comparison of positive with negative feedback allows a differential assessment of positive versus negative prediction errors. Areas activated significantly more by positive than negative feedback included bilaterally the nucleus accumbens, the body of the caudate and paracentral areas (BA 4/6) as well as the right parahippocampus and a medial orbitofrontal focus (BA 11). Areas activated significantly more by negative than by positive feedback included the rostral cingulate zone (BA 32/8) and bilaterally the anterior insula as well as a locus in the right middle temporal gyrus (BA 21) (see Table 3.2 and Figure 3.5 F).

Effects of training

As the behavioral data indicated significant effects for the main effect of *block* (first / second block of the task) and the interaction between *block* and *condition* (cognitive feedback / monetary reward), we ran the first level fMRI analysis again with separate regressors for the first vs. second block to check

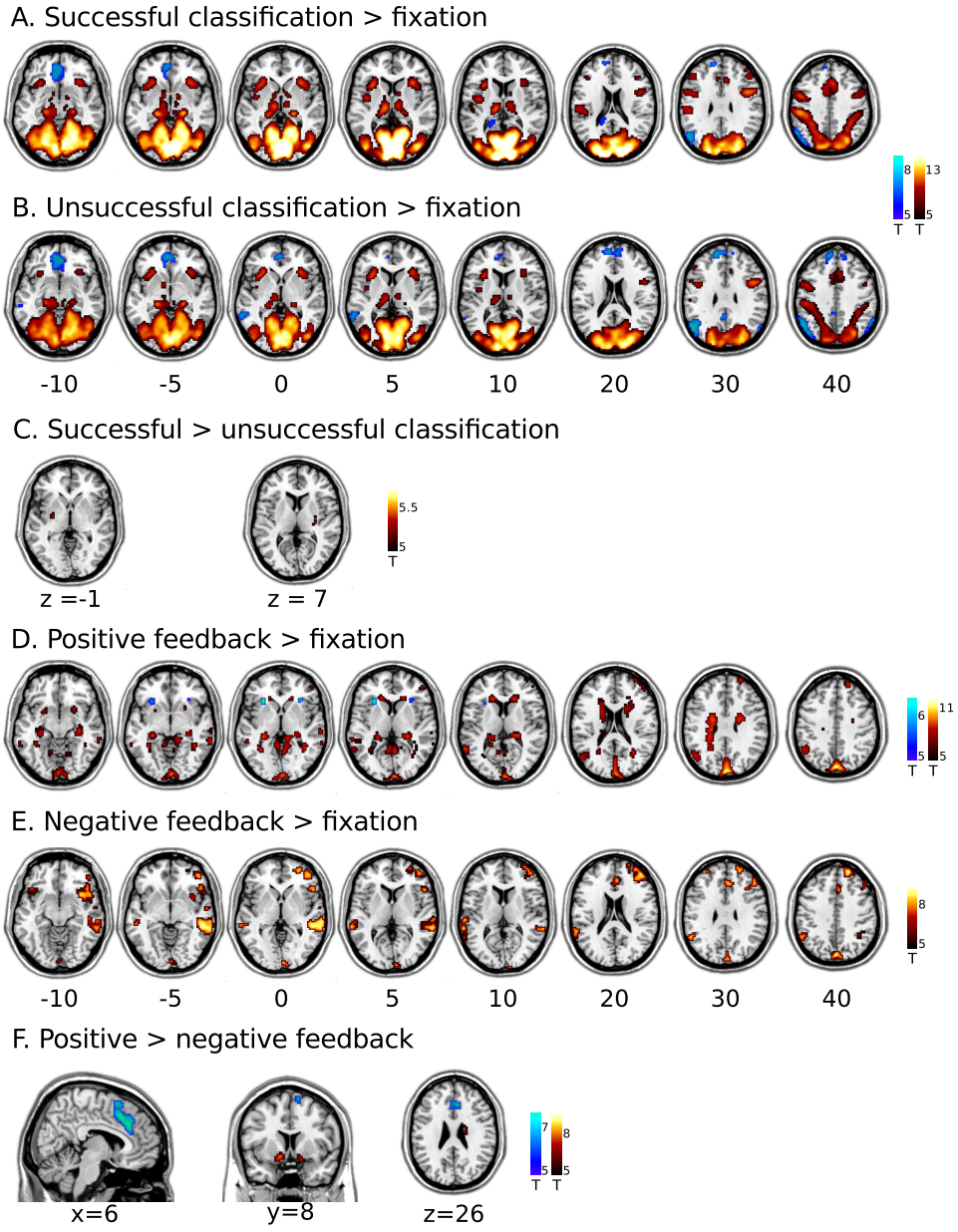

Figure 3.5. Further fMRI results. **A**,**B** and **D**,**E**. Activations (yellow to red) and deactivations (white to blue) for contrasts against fixation. **A.** Successful categorization minus fixation. Activations are observed in occipital and parietal cortices as well as in subcortical areas. **B.** Unsuccessful categorization minus fixation. Activations are mainly observed in occipital and parietal cortices. **C.** Activations in the contrast of successful minus unsuccessful categorization. No voxel showed higher activations for unsuccessful categorization, while bilateral clusters of higher activation during successful categorization were observed in the putamen. **D.** Positive feedback minus fixation. Both the caudate nuclei and the hippocampi are activated. **E.** Negative feedback minus fixation. Activations include the rostral cingulate zone and right prefrontal areas. **F.** Activations in the contrast of positive minus negative feedback. Voxels that were more activated during the processing of negative feedback include the RCZ and anterior insula, while voxels more activated during the processing of positive feedback are observed in the nucleus accumbens and right caudate body. All maps are thresholded $p_{FWE} < .05$, the left hemisphere is presented at the left.

Table 3.2. *Areas of activation when comparing successful and unsuccessful classification and positive and negative feedback*

					MNI		
Region	L/R	BA	k	Max. T	x	y	z
Successful classification - unsuccessful classification							
Putamen	l		10	5.35	-27	-6	-3
	r		6	5.32	30	-12	12
Posterior parietal cortex	l	40	7	4.89	-54	-54	39
Unsuccessful classification - successful classification							
-							
Positive - negative feedback							
Nucleus Accumbens	l		21	8.05	-12	6	-12
	r		7	6.74	15	-6	-15
Caudate body	r		9	5.54	21	-12	27
Parahippocampal gyrus	r		8	5.37	36	-42	-6
Medial orbitofrontal cortex	m	9/10/11	13	5.80	-3	33	69
	m		10	5.39	0	51	-6
Precentral gyrus	l	4/6	11	5.57	-15	-33	69
Negative - positive feedback							
Dorsal anterior cingulate	m	32/8	262	7.09	6	21	42
Anterior insula	r		121	7.15	33	21	-15
	l		70	6.92	-33	21	-15
Middle temporal gyrus	r	21	8	5.67	54	-36	-3

Note. Included regions exceeded an extent threshold of 5 contiguous voxels and a $p < .05$ (FWE-corrected). For each region, the voxel with the maximum T value is described (Max. T). The voxel coordinates refer to the Montreal Neurological Institute (MNI) template. The hemisphere (L/R) of the clusters is indicated (left (l), right (r), medial (m)) and corresponding anatomical labels (Region), Brodmann areas (BA) and cluster sizes (k, 3x3x3mm voxels) are listed.

for neuronal substrates of the behavioral effect. No training effect or *block* × *condition* interaction was observed in the functional data.

3.4. Discussion

3.4.1. Differential activations during categorization

We compared monetary reward with cognitive feedback in information-integration learning. The anticipation of monetary reward led to higher activation than the anticipation of cognitive feedback in a single structure, the nucleus accumbens. Activation in the nucleus accumbens has previously been shown to increase with both reward magnitude and reward probability (Abler et al., 2006; Knutson et al., 2001) during reward expectation and was therefore suggested to code for expected reward value, which is defined as the product of these (Knutson et al., 2005). As in our experiment error rates between the rewarded and cognitive feedback task did not differ, probability of reward was constant across the tasks and the differential activation in the nucleus accumbens is likely to represent an effect of reward magnitude.

Previous results indicate that subcortical dopaminergic structures are implicated in implicit but not explicit category learning (Nomura & Reber, 2008). Although in the current study the task was designed so that optimal performance is only possible when information from both stimulus dimensions is integrated predecisionally, this cannot ensure that participants did not employ explicit rules nevertheless. Therefore a series of models describing the location of responses in the two-dimensional stimulus space was fit to each participant's behavioral data (Maddox & Ashby, 1993). The anticipation of monetary reward as compared to cognitive feedback led to higher activation in the nucleus accumbens only in the group of participants whose behavioral data were better fit by the information-integration model than by the rule-based models, a result that further underscores the dissociation between verbal and implicit systems for category learning.

We did not observe any further significant differences during reward expectation between the two task versions. This was not due to a lack of activation in the dopaminergic pathways and their target structures. Activations against baseline within the head of the caudate nucleus, the pallidum and midbrain were observed during categorization. Also, differential activations were present in a dopaminergic target structure, the putamen. It was significantly more activated during successful as compared to unsuccessful categorization. This fits in well with previous studies on classification learning which reported activation within the putamen (Cincotta & Seger, 2007) that correlated with accuracy and increased with training (Seger & Cincotta, 2005). With its connections to premotor areas, the putamen has been suggested to be central to action selection (Seger, 2008) and to be implicated in the skilled performance of a task (Poldrack et al., 2005).

3.4.2. Differential activations during feedback processing

Next to reward expectation we also investigated activation related to another central aspect of learning, the prediction error. Given their experience with the task, the present experiment participants had to expect positive feedback / reward for about 80% of the trials. Therefore, in both task versions for positive feedback / reward a small positive prediction error can be assumed, and a larger negative reward prediction in response to negative feedback. Two areas previously implicated in processing reward prediction errors were differentially activated during the processing of negative and positive feedback: the RCZ was more active during the processing of negative feedback, while the nucleus accumbens was more active during the processing of positive feedback. The finding of RCZ activation in response to negative feedback is in accordance with a large body of research (for an overview see e.g. Ridderinkhof, Ullsperger, et al. (2004)). This activation is often interpreted as reflecting an transmission of the prediction error signal originating in the mesencephalic dopaminergic system (Holroyd & Coles, 2002),

which in turn signals other brain areas the increased need for control to induce behavioral adjustments and thereby maximize performance (Ridderinkhof, Wildenberg, et al., 2004). Activation within the nucleus accumbens has previously been shown to reflect the positive prediction error both in dependence on the probability (Abler et al., 2006) and on the magnitude (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001) of reward. The observed activations both within the RCZ and the nucleus accumbens only reflected reward valence and were not further modulated by the different types of reward presented. Concerning the RCZ activation, this is in line with studies on the ERN / FRN, an event related EEG component, which is thought to be generated in the RCZ (Debener et al., 2005), and has been shown to be only modulated by feedback valence but not by the magnitude of the (not received) reward (Yeung & Sanfey, 2004; Hajcak, Moser, Holroyd, & Simons, 2006). Previous experiments on reward in humans showed increasing activation in the ventral striatum during the anticipation of increasing monetary rewards (Knutson et al., 2001; Tobler et al., 2007), while effects of reward magnitude on the processing of actual reward are less clear. In studies where the height of the reward was subject to a prediction error, a positive relationship to the magnitude of outcome was observed in the ventral striatum (Breiter et al., 2001), while Delgado, Locke, Stenger, and Fiez (2003) only observed dorsal striatal responses.

During the processing of feedback we also observed differential activation in the body of the caudate nucleus. This structure is part of the visual cortico-striatal loop and receives both highly compressed input from visual cortices as well as a dopaminergic signal. Therefore the caudate body and tail are regarded as the central structures for the establishment of stimulus-response contingencies in information-integration category learning (Ashby & Maddox, 2005; Nomura & Reber, 2008; Seger, 2008). This interpretation is supported by our finding that the caudate body is significantly more activated during the processing of positive than during the processing of negative feedback. Again, the activation within the body of the caudate nucleus was not significantly modulated by the type of reward, i.e. if monetary reward or cognitive feedback was presented.

Additionally, we observed activation in the right posterior parahippocampus that was stronger for positive as compared to negative feedback. No differential medial temporal lobe activations were observed for successful as compared to unsuccessful task performance. This result is in line with Seger and Cincotta (2005), who reported hippocampal and parahippocampal activation during implicit category learning associated with the processing of positive feedback, but not with correct classification.

In summary, we observed several differential activations in dopaminergic projection areas during the processing of feedback, however none of those was modulated by our reward manipulation. The reward magnitude within a given condition of our paradigm was fixed, leaving only the question whether reward on a given trial will be delivered or not.

3.4.3. Commonalities of reward and cognitive feedback-based learning

The idea that reward and cognitive feedback based information-integration learning share similar functional substrates first came up a decade ago (Ashby et al., 1998). It has received considerable support since then (Ashby & Maddox, 2005; Nomura & Reber, 2008), but no study has directly compared the processes within a single fMRI experiment. In this comparison, we observed several differential activations within dopaminergic projection areas in the striatum during information-integration learning, including activation in the putamen for successful as compared to unsuccessful categorization, and in the nucleus accumbens and in the body of the caudate nucleus for positive as compared to negative feedback. However, none of these activations was significantly modulated by the type of feedback, whether cognitive or monetary. The only difference between monetary reward and cognitive feedback we observed was a quantitative effect within the nucleus accumbens during categorization, i.e. while participants anticipated the reward. This result is in line with the results of previous experiments on monetary reward expectation, which show that activation in the nucleus accumbens during reward anticipation increases with the magnitude of potential gains (Knutson et al., 2001, 2005). Taken together with our results this may, on first sight, indicate that cognitive feedback and monetary reward are processed very similarly with the subjective incentive magnitude as only difference. This interpretation

would predict that, when parametrically varying the magnitude of the monetary reward, the anticipation of low monetary gains elicits a very similar reaction as the anticipation of cognitive feedback. However, when correlating the individuals' motivational states with the BOLD signal change, we observed that the activation within the nucleus accumbens increased with the participant's perceived competence during the expectation of cognitive feedback and with the subjective pressure / tension during the expectation of monetary reward. Perceived competence is assumed to be a predictor of intrinsic motivation, while pressure and tension are predictors of extrinsic motivation (Deci, Eghrari, Patrick, & Leone, 1994; Ryan & Deci, 2000). Observing this distinct pattern within the same subjects suggests that, while the accumbens parametrically codes for the incentive value of a potential reward, it does so distinctly for different kinds of motivation. Following this line of argumentation, it is expected that even small monetary rewards may alter motivational processes. Building on the present results, it is an interesting topic for further studies whether motivational changes are associated with different neural activation patterns within dopaminergic structures.

3.4.4. Summary

Cognitive feedback and monetary reward activate dopaminergic structures in a very similar way during information-integration category learning, which supports the assumption that forms of learning which depend on response contingent feedback rely on similar neuronal substrates as reward learning (Ashby & Maddox, 2005; Nomura & Reber, 2008; Seger, 2008). Differential activations in dopaminergic target areas which were not significantly affected by altering the type of reward were observed in the putamen for successful as compared to unsuccessful categorization, and in the nucleus accumbens and in the body of the caudate nucleus for positive as compared to negative feedback. Only activation in the nucleus accumbens during reward expectation differed in dependence on the type of expected outcome information, and this activation was predicted by individual motivational states. The present results therefore indicate that the nucleus accumbens selectively responds to the positive incentive properties of an expected reward in dependence on the specific type of reward.

4 Experiment 2: Striatal activations signal prediction errors on confidence in the absence of external feedback

The results of this experiment were first published in: Daniel, R., & Pollmann, S. (2012). Striatal activations signal prediction errors on confidence in the absence of external feedback. *NeuroImage*, *59*(4), 3457–3467.

4.1. Introduction

Choices are often guided by past experience about which behavior in a given situation provided the most favorable outcome. Computational theories suggest that during reinforcement learning associative links between stimuli and outcomes are formed and adjusted after each trial to minimize future errors in prediction of the outcome (Rescorla & Wagner, 1972; Sutton, 1988). Using fMRI, variables from these models have been shown to be reflected in the activation of dopaminergic midbrain nuclei and their projection sites also in the human brain. Especially the striatum is activated by a wide range of reinforcing stimuli including juice, odors, money and beauty (see e.g. Abler et al. (2006); Aharon et al. (2001); Gottfried et al. (2002); McClure et al. (2003)), but it also responds to information about the correctness of an answer (Aron et al., 2004; Daniel & Pollmann, 2010; Rodriguez et al., 2006). Dopaminergic neurons which project, among other regions, to the striatum, however do not to respond to reward per se. Their activation pattern does not change in response to well-predicted rewards. Rather, they show phasically increased activation when an unexpected reward occurs (a positive prediction error), and a phasic decrease in activation when an expected reward is omitted (a negative prediction error) (Schultz, 2007) (see also Section 1.4).

Models of reinforcement learning are extremely useful for explaining the gradual learning of stimulus-response habits, yet learning can occur in various ways. One of the most studied forms of human learning is category learning, and research its neural bases has mainly focused on paradigms that require the viewing of a stimulus and a categorization response from the observer, after which corrective feedback is provided (Ashby & Maddox, 2005). For this paradigm, which is very similar to instrumental reinforcement learning, the central role of the striatum is well supported. Yet, although feedback-based learning is the most effective training for complex category structures, these can also be learned by observational learning where observers are merely exposed to a stimulus and are told about its associated category label (Ashby & Maddox, 2002; Cincotta & Seger, 2007; Shohamy et al., 2004). Because in ecologically valid situations feedback might not be immediately available after each decision (Hogarth, 2006), the question arises which mechanisms support learning in the absence of feedback. At least two findings indicate that internal signals on potential outcome are present during learning and might therefore be

used to promote learning in the absence of external outcome information: humans have been shown to code the most probable outcome of non-feedback trials into memory (Henriksson, Elwin, & Juslin, 2010), and are capable of generating internal error signals that lead to activations in the RCZ, a region connected to the dopaminergic midbrain (Ridderinkhof, Ullsperger, et al., 2004). Thus, in the absence of feedback, knowledge of the structure of the environment may be used to dynamically self-generate an internal signal on probable outcome, which then drives dopaminergic responses in a similar way as external outcome information. Such a mechanism would support adaptation to the environment whenever the internal model is correct or a sufficient proportion of decisions receives feedback to update it.

The present experiment set out to test the hypothesis that dopaminergic target regions are activated by internal signals and therefore might contribute to learning in the absence of feedback. To this end we employed a standard information-integration category learning task. In feedback-based learning this task activates dopaminergic projection sites (Daniel & Pollmann, 2010; Nomura et al., 2007) including the striatum and RCZ, and patients suffering from diseases impairing striatal functioning show impaired performance (Filoteo et al., 2001). Eighteen participants underwent extensive observational training on stimulus category membership in four sessions within one week. Between the observational blocks they were asked to categorize test stimuli without feedback. FMRI data was acquired during the first and last session. Dopaminergic nuclei have also been shown to respond to stimulus novelty (Bunzeck & Düzel, 2006). Thus, decreasing activations with learning might be due to the increasing familiarity of the stimuli without any relation to feedback processes. We therefore expected striatal activation during the observation of stimuli to decrease with increasing exposure to the stimuli. However, if participants are able to internally generate signals on potential errors during the test blocks, comparing correct to incorrect answers should lead to similar activation differences as comparing positive to negative feedback in feedback-based learning (see Daniel and Pollmann (2010) and Chapter 3 for results on the same task). For a further examination of the underlying mechanisms we assessed the putative internal signals on probable outcome by explicitly asking participants to indicate after each trial how confident they were about their last decision. Within this framework the subjective confidence can be interpreted as a measure of internal processes which are either rewarding themselves, since they indicate correct task performance, or a signal for the possibility of future reward. As dopaminergic neurons do not respond to the absolute magnitude of reward but to deviations from the expected reward (Schultz et al., 1997), we inserted the confidence rating as outcome measure into a standard reinforcement learning model. This allowed us to estimate the expected outcome, i.e. the expected level of confidence, on each trial. If the striatum is activated during observational category learning due to internal signals on the correctness of the outcome, striatal activations are expected follow the prediction error on confidence, i.e. the difference of the expected and the actual confidence.

4.2. Methods

4.2.1. Participants

Eight male and ten female subjects with an average age of 24.4 years [range = 20-29; SE = .5] participated in the experiment. All of them received an allowance of 36 €.

4.2.2. Stimuli

To ensure comparability of the results the same stimuli as in Experiment 1 (Chapter 3) were utilized. However, as only one task was performed by each participant, half of them were presented with the circle stimuli, the other half with the parallel line stimuli, and only the category structure with a positive slope of the decision bound was used. The category structure is depicted along with example stimuli in Figure 4.1 and all parameter values are listed in Table 4.1.

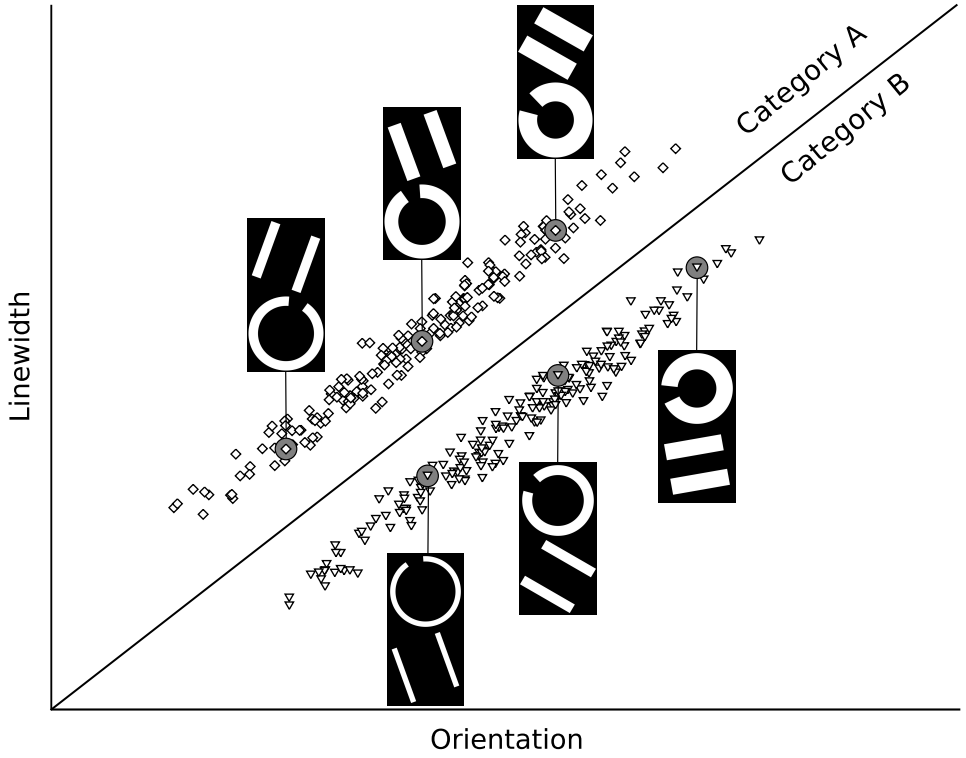

Figure 4.1. *Category structure and sample stimuli.* Each square denotes the orientation and linewidth of a single stimulus from Category A, a triangle those of a stimulus from Category B. The optimal decision bound is represented by a line, all stimuli above it are members of Category A, and all stimuli below it are members of Category B. To establish independence of the results from specific stimulus properties half of the participants were presented with circular stimuli, the other half with lines. For both types of stimuli used in the experiment examples of three stimuli from each category are shown.

4.2.3. Procedure

All participants performed four sessions on separate days, during the first and last of which fMRI data was acquired. The fMRI sessions were conducted spaced at an interval of one week. Between them each participant underwent two training sessions during which only behavioral data was acquired. Within each session two tasks, an observational and a test task, were presented in blocks in an alternating manner. Each session started with a block of observational trials in which first a category label was presented for 500 ms and was directly followed by a stimulus. The next trial began after an intertrial interval (ITI) that was sampled from an exponential distribution with a mean of 6 s [range = 1 - 12 s] in the fMRI sessions and a mean of 3.5 s [range = 1 - 6 s] in the training sessions. Participants were instructed to closely observe the provided examples and attend to both linewidth and orientation to learn which stimuli belong to which category. When learning by observational training, participants are often required to press the button associated with the stimulus to ensure that sufficient attention is paid and that a motor association between stimuli and responses is build. Actively making right responses could induce associative learning, which is a possible confound we wanted to eliminate. Also, motor associations during observational information-integration learning were shown to not significantly influence accuracy rates or the decision bound employed by participants in test blocks (Ashby & Maddox, 2002). This result was confirmed by a pilot study in 24 participants using our stimulus material (Westphal (2009); bachelor's thesis co-supervised by the author of this thesis, see also Appendix B.1). Therefore there was no further behavioral task in the observational blocks. Each observational block was followed by a block

Table 4.1. *Category distribution parameters*

	μ_x	μ_y	σ_x	σ_y	cov_{xy}
Category A	500	700	44100	44100	43500
Category B	700	500	44100	44100	43500

μ = mean for each dimension; σ = variance for each dimension; *cov* = covariance between dimensions.

of test trials where first a stimulus was presented for two seconds. During this period, participants were required to indicate category membership of the presented stimulus as soon as they were reasonably confident about their decision by pressing one of two response keys. After a variable delay sampled from an exponential distribution [M = 2 s; range = .5 - 6 s], the numbers 1 to 5 appeared on the screen for two seconds. Participants were instructed to press one of five buttons to indicate how confident they were about their last decision on a scale from 1 (very unconfident) to 5 (very confident). An additional button was provided to indicate that they had accidentally given a wrong response during the last trial. When the participants used this button, the trial was excluded from further analysis to prevent error-related activity caused by speeded responding from influencing the results. After an ITI that was sampled from an exponential distribution [fMRI session: M = 6 s, range = 1 - 12 s; training session: M = 3.5 s, range = .5 - 6 s] the next trial was presented. During the fMRI sessions, two blocks of observation alternated with two blocks of testing with 50 trials in each block. In the training sessions, four blocks of observation, each with 50 trials, alternated with four blocks of testing composed of 25 trials, so that the overall amount of test trials was the same in fMRI and training sessions, however participants observed twice as many stimuli in the training sessions. A visualization of both an observational and a test trial as well as of the structure of the whole experiment is provided in Figure 4.2. Please note that participants never received any feedback during the whole experiment.

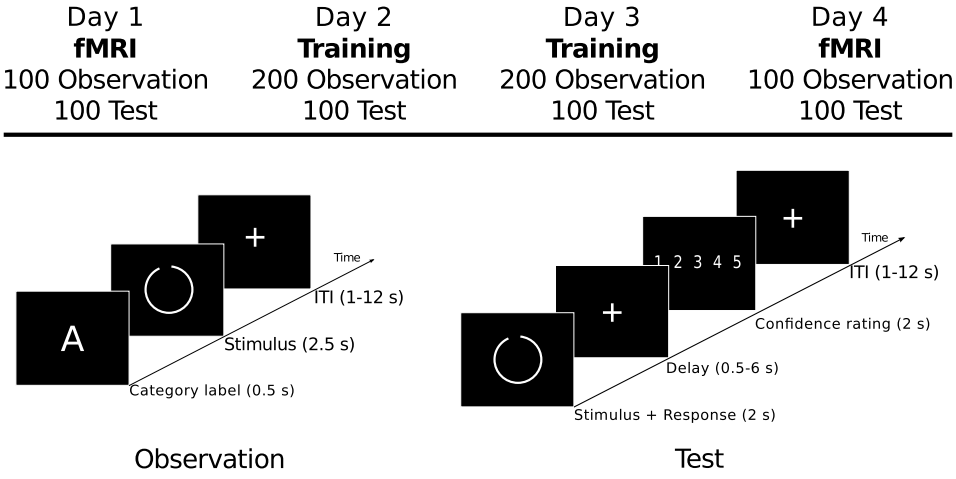

Figure 4.2. *Procedure.* Each participant performed four sessions on different days, the first and last of which were fMRI sessions. FMRI sessions included two blocks of 50 observational trials and test trials, the training sessions included four blocks of 50 observational trials and 25 test trials. Observational trials started with the presentation of category label directly followed by a stimulus. Test trials started with a stimulus to which participants were requested to respond with a decision about its category membership. After a variable delay they rated how confident they were about their last decision. No feedback was provided during the whole experiment.

4.2.4. Model-based analysis of individual decision bounds

Participants learning an information-integration categorization task by observation often do not employ the optimal decision bound depicted in Figure 4.1, but rather a suboptimal strategy relying on explicit verbal rules (Ashby & Maddox, 2002). To ensure that participants were able to acquire the experimenter-defined optimal decision bound, model-fitting procedures as described in Section 2.3 were applied to each participant's response data separately for each session.

4.2.5. Reinforcement learning model

Behavioral data from all participants was modeled using a simple reinforcement learning model (Rescorla & Wagner, 1972; Watkins, 1989). The model consists of two input and two output nodes and assumes that participants update the value Q of a choice c_t (= respond A or B) given a stimulus on each trial t (Gluck & Bower, 1988b, 1988a; Rodriguez et al., 2006). Values Q were initialized to zero and updated after each trial according to an error-correction rule with a free learning rate parameter α:

$$Q_{t+1}(c_t) = Q_t(c_t) + \alpha \cdot \delta_t$$

δ_t represents the prediction error and is computed as

$$\delta_t = r_t - Q_t(c_t)$$

with r_t being the received reward at trial t. On test trials, the value for reward that was input into the model was the reported confidence level, which was linearly scaled to lie between 0 and 1. On observational trials the confidence (and therefore reward) was assumed to be 1. Weights from the input nodes (stimulus from category A or B) to the output nodes (answer "A" or "B") were updated on each trial according to the stimuli presented and choices made. As on test trials the choice could be wrong with a high confidence, also weights on the path from the input node stimulus A to the output node response B and vice versa could be strengthened. Please note that this approach assumes that participants have a representation of which category the presented stimulus is from, which might only be true at late stages of the learning process. However, the present experiment did not aim at developing a new model for learning in the presence of uncertainty, but rather compare the neural correlates of internal signals on correctness with those previously observed during reward-based learning. Therefore an approach was chosen which ensures maximum comparability to previous analyses of reward-based learning, where each class of stimuli is associated with a distinct value, rather than assigning a common value to all stimuli presented in the experiment.

To estimate the free learning rate parameter, it was assumed that participants choose probabilistically from a *soft-max* distribution (Daw, 2011):

$$P(c_t = A|Q_t(A), Q_t(B)) = \frac{exp(\beta \cdot Q_t(A))}{exp(\beta \cdot Q_t(A)) + exp(\beta \cdot Q_t(B))}$$

The free parameter β is the inverse temperature parameter which can be interpreted as reflecting the randomness of the individual's choices. It cannot be assumed that observational learning follows identical rules as learning from the putative internal signal on correctness, therefore α was allowed to be different for observational and test trials. As the models were fit to generate model-based parametric modulators for the imaging analysis based on the prediction error rather than comparing the fit of different models, learning rate parameters were constrained to be equal across participants to regularize the individual estimates which are often noisy (Daw et al., 2006), while β was estimated individually for each participant to account for interindividual differences. All free parameters were estimated simultaneously using maximum likelihood estimation with the optimization package of the SciPy toolbox for Python (Jones, Oliphant, & Peterson, 2001). The inverse temperature was estimated with a mean of 4.112 (SD = 1.020)

across participants, the estimated learning rate was .021 during test trials and .005 during observational trials.

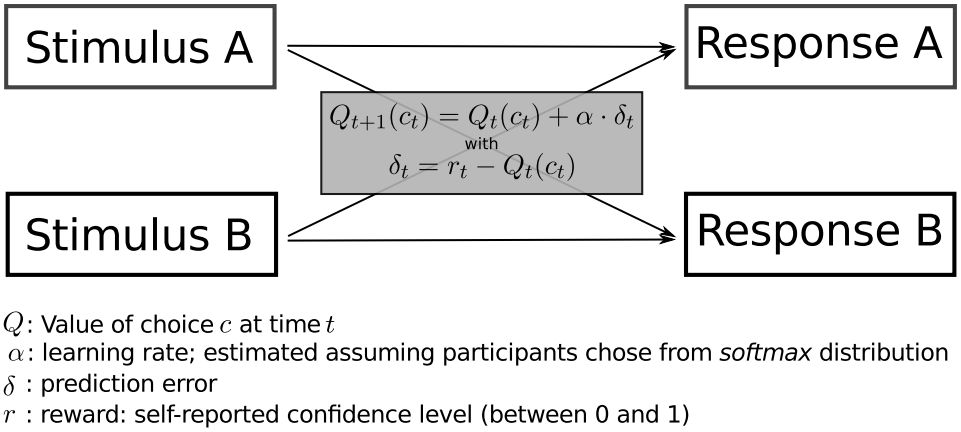

Figure 4.3. *Calculation of the prediction error on confidence.* The arrows between stimuli and response represent the value Q of choosing a certain response (pressing button A or B) given (a) the present stimulus (A or B) and (b) the time or trial number of the experiment. The model was calculated individually for each participant. At the first trial of the experiment all values are zero and get updated after each choice depending on how confident the participant reported to be in that choice. From this model δ_t, the prediction error on confidence for each specific trial, was derived. It is high when the participant reported to be more confident than could be expected from the current value of the chosen option. The prediction errors obtained by this calculation were used in the fMRI analysis to determine which areas within the striatum and midbrain show higher activation when the participant is more confident than expected.

4.2.6. FMRI Image Acquisition and Processing

The parameters for image acquisition are detailed in Section 2.4. Functional images were acquired in a single run of approximately 36 minutes for each imaging session and processed with the statistical parametric mapping software SPM5 (Wellcome Department of Cognitive Neurology, London, UK). After discarding the first five images, slice timing correction with the first slice as reference and realignment to the mean of all images was performed. The estimated movement parameters were saved and later included as regressors of no interest in all statistical analyses. Images were normalized to MNI space (Evans et al., 1993) using the standard EPI template of SPM5. A temporal highpass filter of 1/128 Hz was applied to remove low frequency confounds and all images were smoothed by a 6 mm FWHM Gaussian kernel. All coordinates are reported in MNI space.

4.2.7. Statistical analysis of the fMRI data

In the first analysis step the three regressors of interest represented the observational task as well as correct and incorrect performance of the categorization task separately for both fMRI sessions. An additional regressor represented the time at which the display prompting participants to give the confidence rating was presented. If the participant failed to respond on any trials an additional regressor representing the missed trials was included. The onset of each event was set to the beginning of the visual stimulation, its duration was set to an epoch with the duration of the visual stimulation (see Figure 4.2) and convolved with a model of the hemodynamic response function as implemented in SPM5. The regressors were fit separately to the data of each participant using the GLM. Data from separate days was treated as different sessions. To test whether the activation in dopaminergic projection sites decreases with stimulus novelty during the observation, the regressors for the observational task from the first and second fMRI session were submitted to a second level random effects analysis with *session*

as a single factor. For the examination of effects of the valence of putative internal signals indicating an outcome, a 2 × 2 random effects ANOVA with the factors *correctness* and *session* was calculated on the regressors for correct and incorrect task performance. Three participants had to be excluded from this analysis because they committed less than two error trials in the trained session. The activation maps of these analyses were thresholded at a level of $p < .05$ (spatial extent of more than two contiguous voxels), and all reported clusters survived correction for multiple comparisons at the whole-brain level using FWE correction (Worsley et al., 1996).

A second GLM analysis was performed to test for the assumption that differential activations during correct and incorrect task performance are attributable to signals following a similar pattern as the prediction error signals during reinforcement learning. The two regressors for correct and incorrect task performance were replaced by a single regressor containing all categorization task events and a linear parametric modulator was added to the regressor. The height of this modulator corresponded to the prediction error of the trial as computed using the reinforcement learning model described in Section 4.2.5. To increase sensitivity, this analysis was only performed in areas of interest. They comprised the caudate nucleus, the putamen and the nucleus accumbens as defined by the Harvard-Oxford subcortical structural atlas implemented in the FSL-FMRIB Software Library (http://www.fmrib.ox.ac.uk). These areas constitute the striatum and are dopaminergic target regions which have previously been reported to show activation changes as predicted by reinforcement learning models (Niv, 2009). Additionally, a sphere of 15 mm radius centered on the midbrain (MNI: x = 0, y = 15, z = 9) was included (Aron et al., 2004), as projections from this area are considered to be the source of striatal activations. The fMRI activation map was thresholded at $p < .001$ and activations were considered significant at a level of $p < .05$ after correction for multiple comparisons at the cluster level.

4.3. Results

4.3.1. Behavioral results

Accuracy measures

To ensure that the observational training led to successful learning and the two task versions were comparable in difficulty, the error rate data of all 18 participants was subjected to a 4×2 ANOVA with repeated measures on *session* and the *type of stimuli* (lines or circles) as between-subjects factor. Results show a significant effect of *session* $[F(2.1, 34.3) = 32.7, p < .001]$, indicating that the error rates decreased from 25.9 % (SD = 11.7) in the first session to 7.3 % (SD = 7.0) in the fourth session. No effect of *type of stimuli* $[F(1, 16) = .98, p = .34]$ or interaction of *session* × *type of stimuli* $[F(2.1, 34.3) = .53, p = .61]$ was observed. To test for possible behavioral learning effects within the test blocks the means of the first and last 20% of trials within each block were compared using multiple T-tests. No significant results were observed at the .05 level (uncorrected) (Figure 4.4).

Correlations between reaction times, confidence ratings, the prediction error and error rates

For the calculation of the prediction error, which was input as a parametric modulator into the analysis of the functional imaging data, it was assumed that confidence ratings can be treated as a measure of internal signals on correctness. For this assumption to be plausible the confidence ratings should correlate negatively with reaction times and error rates. For each participant Pearson's correlations between reaction times and prediction errors as well as confidence ratings across all days were calculated. The mean of these values was determined using Fisher's z-transformation and tested against zero. The backtransformed mean correlation between reaction time and confidence rating was -.47 $[T(17) = -13.10, p < .001]$. Please note that there was no correlation between reaction times and prediction errors $[r = -.005; T(17) = -.10, p = .92]$. This is to be expected since the same level of confidence, associated with a similar reaction time, can be higher than expected at the beginning of

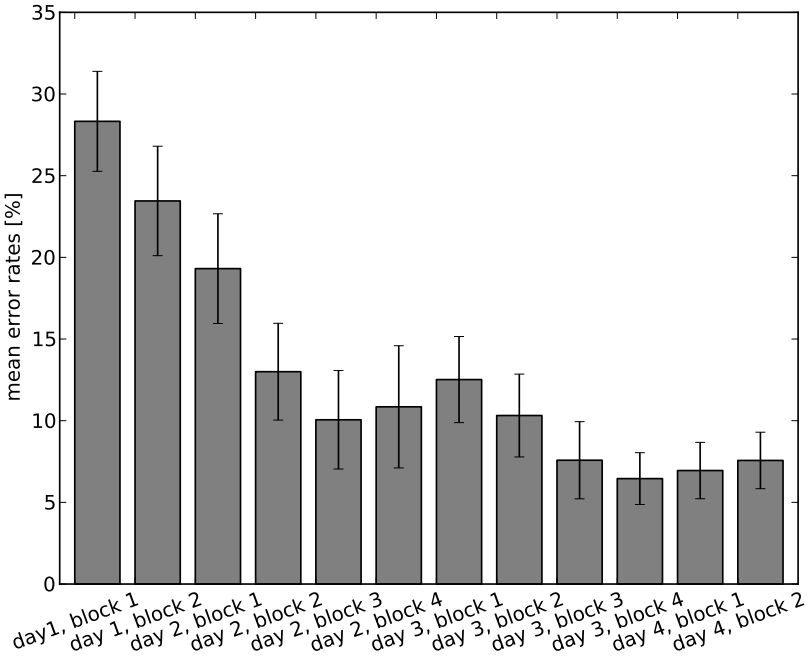

Figure 4.4. *Behavioral training effect.* Mean error rates for each day and block are depicted, errorbars represent the standard error of means.

the experiment and lower than expected at the end of the experiment. This means that correlations between the fMRI data and the prediction error cannot be caused by reaction time differences. To determine correlations with the error rates, they were calculated separately for each block. Reaction times and confidence ratings were averaged blockwise analogously. Error rates correlated significantly with both mean confidence ratings $[r = -.78; T(17) = -10.80, p < .001]$ and mean reaction times $[r = .53; T(17) = 5.05, p < .001]$.

Model-based analysis of individual decision bounds

In the first session the data of three of the 18 participants was best fit by a unidimensional model, the data of eleven participants was best fit by a conjunction model and only five participants employed the optimal rule. From the third session on this relationship was reversed and all but one of the participants employed the optimal rule, indicating successful learning of the experimenter-defined category structure. The numbers and percentages of participants best fit by each model are depicted in Figure 4.5, BIC values are listed in Appendix B.2.

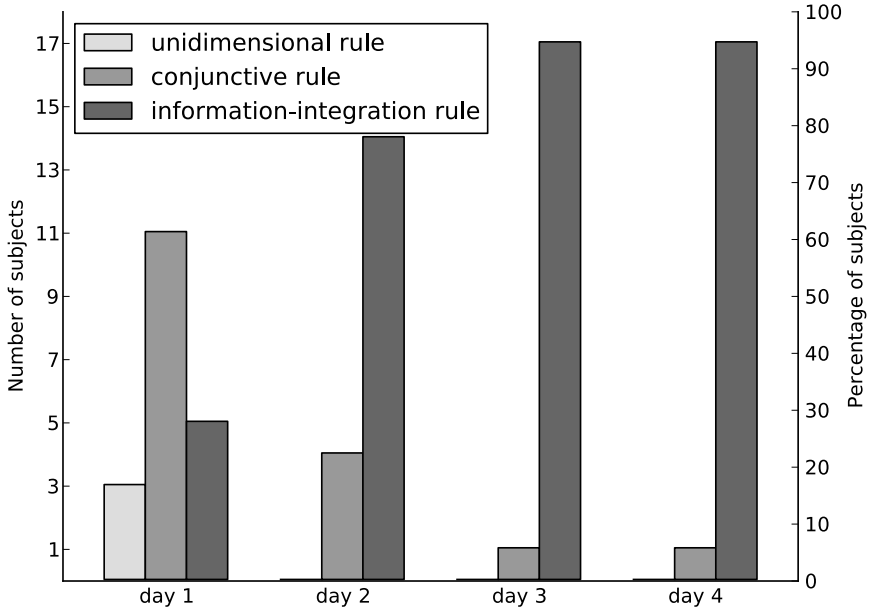

Figure 4.5. *Results of the model-based analyses.* The number and percentage of participants best fit by models assuming the participant used a unidimensional, conjunctive or information-integration decision bound. All results are corrected for the number of free parameters. From the third session on all but one participant used the ideal information-integration decision bound, indicating that the complex categorization task was learned after a sufficient number of observational training trials in the absence of external feedback.

Table 4.2. *Differential fMRI activations for right and wrong answers*

					MNI		
Region	L/R	BA	k	Max. T	x	y	z
Untrained categorization task: right > wrong							
Nucleus Accumbens	l		4	5.95	-6	6	-12
Putamen	l		2	5.46	-27	0	6
	r		19	6.47	30	3	9
	r		13	5.87	27	-15	6
Caudate Body	l		2	5.49	-21	-12	21
Anterior Cingulate Cortex	l	32	2	5.5	-12	39	12
Posterior Parahippocampus	l		9	6.85	-42	-36	-6
Precentral Gyrus	r	6	4	5.72	9	-21	51
	r	6	5	5.52	24	-21	60
Superior Medial Gyrus	l	10	2	5.63	-9	60	21
Trained categorization task: wrong > right							
Dorsal medial prefrontal cortex	b	6/8/32/24	169	7.58	-3	18	51
		6	10	6.32	-12	9	57
Anterior Insula	l		85	7.27	-30	21	-9
	r		59	7.14	39	15	-3
Midbrain	b		10	5.76	3	-24	-30

l=left, r=right, b=bilateral, BA=Brodmann area, k=cluster size in voxels (3x3x3 mm), max. T=maximum T-value

4.3.2. Functional imaging results

Comparing right and wrong answers in the categorization task

In the naïve session activations were higher for right than wrong answers mainly bilaterally in the putamen, but clusters in the caudate body, nucleus accumbens and posterior parahippocampal gyrus were also observed. In this session no clusters showed higher activation for wrong than for right answers. In the trained session this, however, was observed in the dorsal medial prefrontal cortex, the anterior insula and the midbrain. Here, no clusters showed higher activations for right than for wrong answers (Table 4.2, Figure 4.6). Despite this interaction between the factors *session* and *correctness* no significant main effect of *session* was observed, indicating that there was no significant increase or decrease in activation for correct or wrong answers between sessions.

Parametric modulation by the prediction error

To estimate the trial-by-trial prediction error we fit a reinforcement learning model to the participants' behavioral choices. The model estimates the expected outcome of each option given the present stimulus for every trial based on the sequence of previous outcomes. Outcome was quantified by the confidence ratings as reported after each trial, and the prediction error as the difference between the expected and the actual outcome (see Section 4.2.5).

Results show that within our areas of interest (striatum, midbrain) the putamen and nucleus accumbens showed linearly increasing activation with an increasing prediction error on confidence. This indicates that activation was higher whenever the participant reported to be more confident on the trial than could be expected from the participants' own history of responses and lower when he or she reported to be less confident (Table 4.3, Figure 4.7).

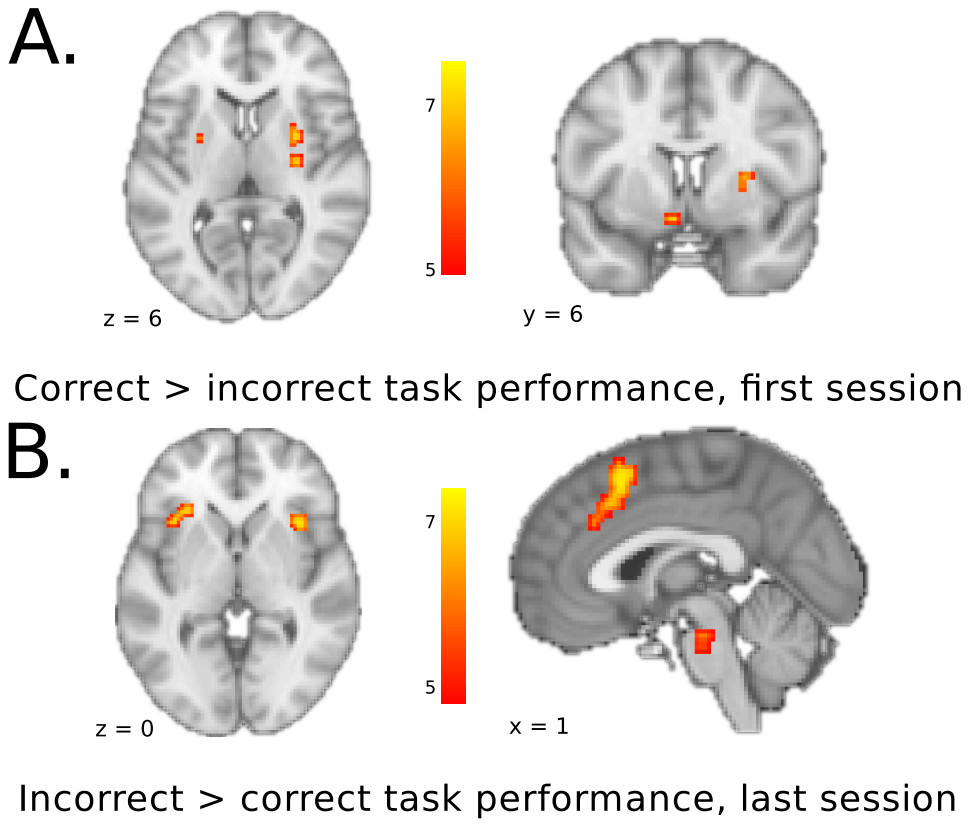

Figure 4.6. FMRI results from the test blocks. **A.** Activations that were higher for right than for wrong answers in the categorization task included bilaterally the putamen and the left nucleus accumbens in the untrained session. No activations that were higher for wrong than for right answers were observed in this session. **B.** Activations that were higher for wrong than for right answers in the categorization task were observed in the midbrain, the posterior medial prefrontal cortex and bilaterally in the anterior insula in the trained session. No activations that were higher for right than for wrong answers were observed in this session. Activation maps are thresholded at $p_{FWE} < .05$; the left hemisphere is presented at the left.

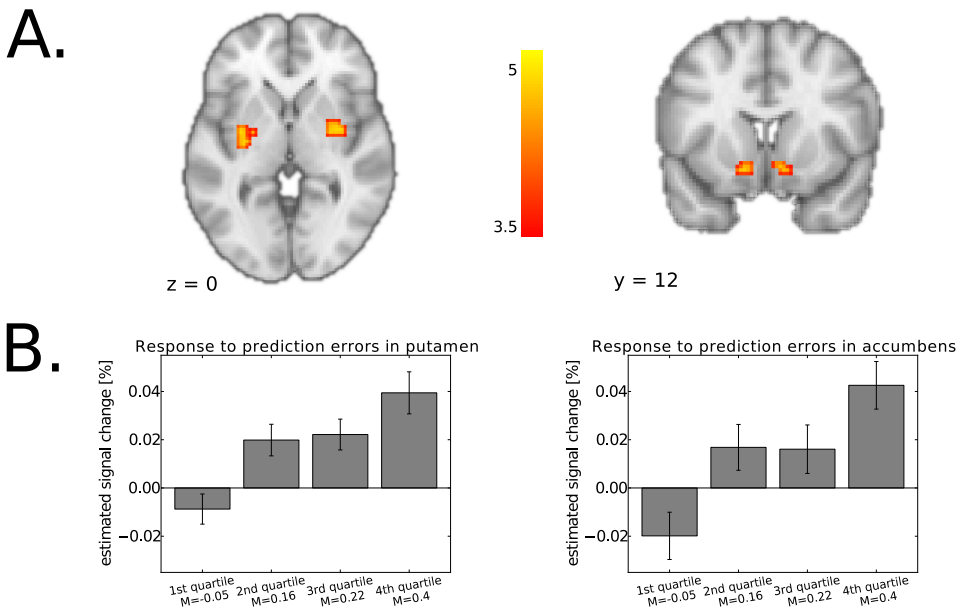

Figure 4.7. *FMRI results on the prediction error.* **A.** Activations of the parametric modulator reflecting the estimated prediction error based on confidence ratings during the categorization task in areas of interest (caudate, nucleus accumbens, putamen, and midbrain). Activations were observed bilaterally the putamen and the nucleus accumbens. All shown clusters are above a cluster level threshold of .05; the left hemisphere is presented at the left. **B.** Visualization of the activation in response to the prediction error. Separate regressors were constructed for each quartile of prediction errors in each session, e.g. the regressor for the first quartile represents the 25 trials with the lowest prediction error of a given participant in a given session. The mean prediction errors for each quartile resulting from this procedure are reported in the axis labels. Negative values indicate that the participant was less confident than expected, positive values indicate that he or she was more confident. Percent signal change for each regressor was estimated using MarsBar (http://marsbar.sourceforge.net). Values represent combined estimates from the peak voxels of both hemispheres; error bars show the standard error of means.

Training effects

Activation decreases between the naïve and the trained session while participants performed the observational task were mainly observed in the inferior temporal / fusiform gyrus, but also clusters in the left lateral prefrontal cortex and the right ventral striatum decreased in activation with training. No clusters were activated more during the trained than during the naïve session in the observational task (Table 4.4, Figure 4.8). When collapsing across correct and incorrect task performance, no effects of training were observed in the categorization task.

Table 4.3. *Regions that show parametric modulation as estimated by the prediction error on confidence in areas of interest (striatum and midbrain)*

					MNI		
Region	L/R	BA	k	Max. T	x	y	z
Nucleus Accumbens	r		13	4.39	12	6	-12
	l		13	4.67	-12	9	-12
Putamen	r		72	4.78	33	-3	9
	l		48	4.41	-30	-12	3

l=left, r=right, BA=Brodmann area, k=cluster size in voxels (3x3x3 mm), max. T=maximum T-value

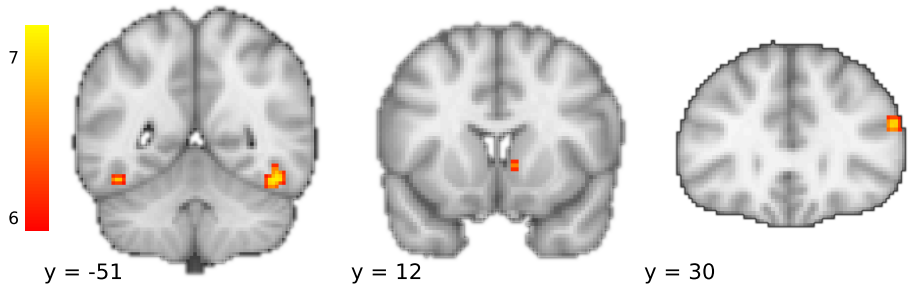

Figure 4.8. FMRI results from the observational blocks. Activations decreasing with training in the observational task from the first to the last session were observed bilaterally in the inferior temporal cortex, in the right ventral striatum and inferior frontal gyrus. No increases in activation were observed in this contrast. Activation maps are thresholded at $p_{FWE} < .05$; the left hemisphere is presented at the left.

Table 4.4. *Regions that show decreased activation with learning in the observational task (observation first session > observation last session)*

					MNI		
Region	L/R	BA	k	Max. T	x	y	z
Nucleus Accumbens/Head of Caudate	r		2	5.95	9	9	0
Inferior Temporal/ Fusiform Gyrus	r	37	21	7.25	45	-51	-18
	r	37	2	6.7	33	-42	-21
	l	37	18	6.27	-45	-57	-12
	l	37	2	6.07	-45	-51	-18
Inferior Frontal Gyrus	r	45	4	6.46	51	30	21
	r	45	2	5.84	48	24	30

l=left, r=right, BA=Brodmann area, k=cluster size in voxels (3x3x3 mm), max. T=maximum T-value

4.4. Discussion

Models of reinforcement learning describing the role of dopaminergic neurons and their striatal projection sites, mainly the ventral striatum, explain the gradual learning of stimulus-response habits. However, recently similar processes have been suggested to participate in more complex tasks. For example, activation within the ventral striatum is modulated by social information, prediction errors in the learning of trust and subjective goals during episodic memory retrieval (Behrens, Hunt, Woolrich, & Rushworth, 2008; Delgado, Frank, & Phelps, 2005; Han, Huettel, Raposo, Adcock, & Dobbins, 2010; King-Casas et al., 2005). Also, distinct but interacting corticostriatal loops have been suggested to underlie both habit learning and goal-directed behavior and jointly contribute to producing complex behavior (Daw & Shohamy, 2008; Valentin, Dickinson, & O'Doherty, 2007).

In everyday life only a small proportion of decisions or judgments taken are followed by immediate feedback (Hogarth, 2006). With such incomplete feedback it is improbable that feedback training is the single prevalent mode of learning in ecologically valid settings. It is not clear whether the dopaminergic system also contributes to learning under circumstances where incomplete or no feedback is provided. We therefore investigated the putative involvement of striatal structures in human category learning in the complete absence of feedback. The results support the idea that these dopaminergic projection sites are activated during observational learning in a similar way as during reinforcement learning, but with respect to self-generated internal signals related to confidence rather than external outcome measures.

4.4.1. Striatal activations during observational learning

Within the test phase of the present experiment, when no external information about category membership was available, the nucleus accumbens and putamen were more activated by correct than incorrect answers. This finding parallels results from feedback-based learning, where a similar activation pattern in striatal areas is present during the processing of feedback (for results on a feedback-based version of the same task see Daniel and Pollmann (2010) and Chapter 3). Previous fMRI results show that activity in the putamen is mainly correlated with stimulus-action-dependent reward prediction (Haruno & Kawato, 2006). This finding fits in well with the known anatomy of the corticostriatal systems, as the putamen is connected with motor areas and therefore in a suitable location to integrate reward expectation with processes that mediate actions to obtain the reward (Haruno & Kawato, 2006; Seger et al., 2010). The ventral striatum on the other hand is connected with orbitofrontal and ventromedial cortex areas known to be involved in the processing of reward and motivation (Haber & Knutson, 2009; Seger et al., 2010).

Ventral striatal activity is consistently observed to be modulated by the reward prediction error as estimated by reinforcement learning models (Brovelli, Laksiri, Nazarian, Meunier, & Boussaoud, 2008; Daw & Doya, 2006; Dayan & Niv, 2008; Rangel, Camerer, & Montague, 2008; Schönberg et al., 2007). In search for the processes underlying the differential activations in response to correctness, a reinforcement learning model was fit to each participant's data. In the present experiment, there were no external reward signals present, however participants reported their confidence in their last decision after each trial as a conscious measure of internally generated signals indicating an assumed success or failure. During categorization, activity in the nucleus accumbens and putamen was significantly modulated by the prediction error as estimated from a standard reinforcement learning model in which the confidence rating was input into the model instead of an external reward signal. Possibly due to the postsynaptic origin of the fMRI signal, no significant modulation of midbrain activity was observed. The knowledge of correct task performance activates the ventral striatum similarly as the receipt of primary or secondary reinforcers, if this information is conveyed by external feedback (Aron et al., 2004; Rodriguez et al., 2006; Daniel & Pollmann, 2010). The current results support the supposition that in the absence of external rewards activation in the striatum can be driven by a prediction error on confidence that utilizes internally generated error signals. Findings from two major lines of research argue for the validity of this conclusion.

First, according to standard reinforcement learning models, only information that is provided by

external outcome feedback is coded into memory. However, cognitive psychology has shown that humans tend to reconstruct missing information using top-down processes. These inferences are often used in the same manner as external information (White & Koehler, 2004) and the ability to distinguish between real and inferred information is limited (Roediger, Watson, McDermott, & Gallo, 2001). In the case of selectively missing outcome information, humans have been shown to actively construct the most likely outcome of the trial and code it into memory (Elwin, Juslin, Olsson, & Enkvist, 2007; Henriksson et al., 2010). At times, this approach can be misleading, so that no behavioral improvements are observable, or it can even lead to deterioration in performance due to the strengthening of false conceptions without any external evidence. However, in many cases a good internal representation of the environment is available and outcome information is scarce or costly to obtain. Under these circumstances the application of previous knowledge can be the most adaptive or only way to promote further learning.

Second, a large body of literature provides evidence for the ability of the brain to generate error signals without external feedback. The RCZ is known not only to respond to unfavorable performance outcomes but also pre-response conflict, decision uncertainty and response errors before external feedback (Ridderinkhof, Ullsperger, et al., 2004). Theoretical accounts for these experimental results have suggested that the performance monitoring system makes use of dopaminergic signals carrying reward prediction errors to determine the outcome of ongoing behavior and indicate the need for the recruitment of adaptive cognitive control mechanisms (Holroyd & Coles, 2002; Ridderinkhof, Ullsperger, et al., 2004; Rushworth & Behrens, 2008). The pharmacokinetics of dopamine in the cortex make it improbable that fast activation changes in the dmPFC are a direct effect of dopaminergic signals from the midbrain rather than a remote effect possibly mediated by the striatum, indicating that cortical dopamine only modulates the effects of other neurotransmitter systems. However, there is a close interaction between performance monitoring reflected in the dmPFC, prediction errors reflected in striatal activation and dopamine release in the midbrain. Moreover, there are indications that performance monitoring might involve top-down modulation of mesencephalic dopaminergic neurons, also via pathways through the striatum. This means that errors detected by the performance monitoring system could lead to decreases in the firing rate of neurons in the dopaminergic midbrain, which in turn would influence the prediction error signal that is generated in the midbrain and conveyed back to higher brain regions (Jocham & Ullsperger, 2009; Ullsperger, 2006, 2010). These assumptions however have to remain a conjecture at the moment, as the exact pathways contributing to performance monitoring in the human brain are only beginning to be understood and the fMRI signal is not sensitive to the neurotransmitter system conveying the signal (see Section 1.3). The performance monitoring system can be informed by different sources about real or potential errors: it can utilize external signals, like punishments or the omission of rewards, but also internal sources, like the detection of competing response tendencies. The error-related negativity (ERN), an event-related potential which has a dipole source in the RCZ (Debener et al., 2005), has a functional interrelation with a very similar event-related potential sharing the same neural sources (Gentsch, Ullsperger, & Ullsperger, 2009), the feedback-related negativity (FRN). When participants are able to generate knowledge about an error without external feedback, an ERN but no FRN is observed, while information that is not available prior to the feedback elicits only the FRN (Heldmann, Rüsseler, & Münte, 2008). This is compatible with the assumption that, whenever an external outcome signal provides additional information, it elicits an error signal. However, internal error signals are generated whenever possible before the feedback arrives.

4.4.2. Activation of the dmPFC and anterior insula in response to errors

In line with prediction error models, after extensive training no higher activation for correct than incorrect task performance in striatal areas was observed. At this stage, the midbrain, posterior medial prefrontal areas and the anterior insula showed higher activation in response to incorrect than correct answers. This result is also paralleled by similar findings in the literature on feedback-based learning where this activation pattern, however, is observed already early in training (Aron et al., 2004; Daniel & Pollmann, 2010; Grinband et al., 2006; Ridderinkhof, Ullsperger, et al., 2004). The dmPFC together with

the anterior insula has been shown to respond not only to response errors, but also to increase linearly in activation with uncertainty (Daniel et al., 2011). Also, the anterior insula has been associated with conscious errors (Craig, 2009). The distinct difference in uncertainty for right and wrong answers as well as the explicit knowledge about it might arise only later in observational learning, when the knowledge of the task structure has reached a certain level of proficiency, while in the presence of feedback it is available as soon as the feedback is delivered. However, as there was no significant activation difference in response to answers between the first and last session, this observation has to be supported by data from further experiments.

4.4.3. Decrease in ventral striatal activation with familiarity

Activation within the nucleus accumbens decreased with training during the observation of stimuli. This decrease cannot be elicited by any form of internal signals relating to confidence since the category label was presented before the stimulus and no categorization response was required. Stimulus novelty is known to activate dopaminergic structures, and might have rewarding properties in itself or be interpreted as a cue for potential reward (Bunzeck & Düzel, 2006; Daw et al., 2006). This mechanism can have positive consequences for the organism, since it leads to exploring new options instead of exploiting those that have already proven to have a high utility. Thus valuable information about more profitable options can be attained, which in turn increases future rewards. Although in the present experiment new category members were presented within each block, they were all drawn from the same distribution, which increases familiarity with the stimuli later in training. While the observed higher activations in response to correct than incorrect categorization cannot be attributed to stimulus novelty, since classification accuracy should be higher for subjectively familiar stimuli than subjectively novel ones, ventral striatal activation during observation early in learning might signal saliency (Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004) and therefore support the allocation of attention to unfamiliar category members.

4.4.4. The use of subjective ratings as confidence measure and the potential origin of the confidence signal

The literature on metacognition provides examples for errors of calibration in subjective ratings on post-decisional confidence in a wide variety of tasks. These range from perceptional discrimination tasks over general knowledge questions to giving eyewitness accounts and unambiguously show that subjective confidence often does not predict task performance (Baranski & Petrusic, 1999; Sporer, Penrod, Read, & Cutler, 1995; Zakay & Tuvia, 1998). Depending on the specific task and its difficulty participants often over- or underestimate the correctness or their decisions. In the present task participants had to infer the structure of a categorization task by observing examples for each category and in subsequent test trials were asked for a confidence rating immediately after their decision. Both error rates and reaction times were highly correlated with these confidence ratings, indicating that for this specific task post-decisional subjective confidence ratings are a sensible measure of the internal processes leading to the decision. Also, the use of a prediction error model greatly reduces the demands on the measure, since only intraindividual deviations from an expectation, which is calculated from the weighted ratings of past trials, are considered. Therefore it is only required that participants respond consistently, predominantly on trials closer together in time. Nevertheless, it has to be noted that subjective confidence ratings strongly build on the human ability for metacognition, and processes that are not accessible to it are not captured by this measure (Allwood, Granhag, & Johansson, 2000).

Observational category learning has been suggested to be supported by visual working memory processes and executive attention (Ashby & Maddox, 2002). In the present experiment we observed a decrease in activation in the lateral prefrontal cortex from untrained to trained observation, an area that is often associated with working memory (D'Esposito, Postle, & Rypma, 2000) and the representation of rules and goals (Miller & Cohen, 2001). Later in training increased activation in the RCZ, an area that

is also activated by response conflict (Ridderinkhof, Ullsperger, et al., 2004), was observed for wrong compared to right answers. Taking these previous findings into account low confidence in the present experiment might be caused by encountering exemplars that do not fit the internal model of category structure held in working memory and therefore induce response conflict. However, as the present study was aimed at establishing the role of internal error signals in learning in the absence of feedback rather than determining their origin, these considerations can only provide hypotheses.

4.4.5. Possible roles of the prediction error on confidence in learning

In the present experiment no reward or cognitive feedback was associated with the stimuli at any stage. Nevertheless, ventral striatal areas that are known to respond to reward prediction errors showed an increased activation whenever internal signals informed the participants that they were more confident about their decision as expected, and a decreased activation when they were less confident about their decision.

This internal signal might serve two distinct but overlapping purposes. In a learning environment where no external feedback is available, any mechanism that assesses current task performance can only be informed by internal signals. Therefore, the internal error signal might substitute for external feedback processes. This interpretation is supported by the fact that the activation patterns observed when comparing correct to incorrect task performance are equivalent to those observed when comparing positive to negative feedback on the same task when external feedback is available (see Daniel and Pollmann (2010) and Chapter 3).

Additionally, reinforcement theories predict that the detection of stimuli which change the probability of receiving reward elicits a reward prediction error (Niv, 2009). Although participants were informed that no reward or feedback would be given within the experiment, it is plausible to assume that acquiring the ability to map the structure of the environment is inherently interpreted as increasing the probability of receiving reward within that environment. Seeing an exemplar that can be categorized with lower confidence than expected can serve as a signal that the model of the environment is wrong and needs to be updated, while higher confidence than expected indicates that the state of the environment has changed towards a higher probability of reward. This interpretation predicts a stronger violation of expectations and therefore more learning whenever confident responses are followed by negative feedback. Such an overcorrection effect has been observed in the domain of explicit knowledge before (Butterfield & Metcalfe, 2006).

4.4.6. Summary

In the absence of external outcome information in an information-integration category learning task higher fMRI activations in the striatum were observed for wrong compared to right answers, and higher activations in the anterior insula and RCZ for wrong compared to right answers. This parallels the pattern of activation observed during the receipt of feedback in a reward-based version on the same task (Daniel & Pollmann, 2010). A potential source of the observed activations lies within the performance-monitoring system which is able to utilize internal information. This interpretation is supported by the finding that in the absence of external feedback striatal activation responds to the prediction error similarly as it does during reinforcement learning, but with respect to signals based on subjective confidence. In addition, ventral striatal activation decreases with stimulus novelty during observation. These results provide a parsimonious account for the neural bases of learning, implying that similar basic mechanisms underlie reinforcement learning and learning in settings where external outcome information is unavailable.

5 Experiment 3: Task-independent categorical representation depends on reward differentially in visual areas and orbitofrontal cortex

5.1. Introduction

Making successful everyday decisions relies on the ability of the brain to extract the current state of the environment from the noisy sensory signals it receives, and to integrate this information with prior experience about the costs and benefits of different behaviors in that state. A fundamental component of the process involves assigning stimuli to categories reflecting their functional relevance (Freedman, 2008). Whenever small changes in input are required to lead to categorically different behavioral responses, an amplified representation of feature values close to the category boundary is functionally advantageous (Roelfsema, Ooyen, & Watanabe, 2010). In accordance with this consideration, after categorization training both effects of between-category separation, i.e. enhanced discriminability between members of different categories, and within-category compression, i.e. reduced discriminability between members of the same category, have been observed (Goldstone, 1994, 1998; Liebenthal, Binder, Spitzer, Possing, & Medler, 2005; Livingston, Andrews, & Harnad, 1998).

For categorical decisions based on small variations of low-level stimuli, behavioral data suggests that perceptual learning within early visual areas underlies this categorical perception effect (Notman, Sowden, & Ozgen, 2005). Electrophysiological and fMRI studies report changes in selectivity and overall levels of activation in the visual processing stream from early occipital areas to the inferior temporal cortex as a result of discrimination training (Gillebert, Op de Beeck, Panis, & Wagemans, 2009; Jiang et al., 2007; Kourtzi, Betts, Sarkheil, & Welchman, 2005; Raiguel, Vogels, Mysore, & Orban, 2006; Sigala & Logothetis, 2002). Additionally, changes induced by categorization training and categorization have been observed in inferior parietal and prefrontal areas and the striatum (Freedman, 2008; Li, Ostwald, Giese, & Kourtzi, 2007; Li, Mayhew, & Kourtzi, 2009, 2011; Op de Beeck, Baker, DiCarlo, & Kanwisher, 2006).

Perceptual learning as the basis of categorical perception is assumed to be based on the interplay of attentional feedback signals originating in higher cortical areas that highlight the chain of neurons between sensory and motor areas responsible for any selected action, and diffuse neuromodulatory signals which gate plasticity in dependence on the outcome of the performed behavior (Roelfsema et al., 2010). Reward is known to influence the representation of stimuli already during early processing (Serences, 2008), and its effects can be observed throughout virtually the whole brain (Vickery et al., 2011). While the central role of the nature of feedback during complex category learning is well established (Ashby & Valentin, 2005; Daniel & Pollmann, 2010), it remains an open question whether rewarding stimuli during categorization training has a lasting influence on stimulus representation in categorization-relevant areas.

To address the issues of lasting influence on representation by categorization training and differences

induced by reinforcers with different motivational value, a three-stage approach exploiting the sensitivity of Representational Similarity Analysis (Kriegeskorte, Mur, & Bandettini, 2008) was employed. Human participants were trained via feedback to categorize stimuli into one of four categories, but only responses with regard to one category were monetarily rewarded. Subsequently, fMRI data was acquired during passive viewing. We first localized areas where the physical similarity structure of the stimuli is represented, which we expected to include mainly early occipital areas, in keeping with the simple structure of our stimuli. Subsequently, areas where the similarity structure of the fMRI activations is better accounted for by the category structure than the physical similarity were identified. In the third step, we investigated in which areas the representation of previously rewarded categories differs from the representation of categories that were not rewarded during training. One central question was whether (1) category learning and (2) the association of particular categories with reward would change the representation of category members already in early visual cortex, which would conflict with a truthful representation of their physical similarity structure. The other central question was whether category structure and reward association would be represented, even in the absence of reward expectation, in dopaminergic projection areas involved in reward learning. We expected that the visual representation was optimized to differentiate between categories and that this differentiation is improved by reward. In contrast, representation in reward-related brain areas was expected to be optimized to distinguish reward-associated from not reward-associated stimuli.

5.2. Methods

5.2.1. Participants

Sixteen subjects with an average age of 23.9 years [range = 20-28; SE = 0.5; 8 males] participated in the experiment. They received a variable payment based on performance for the training session and a fixed allowance of 12 € for the fMRI session.

5.2.2. Stimuli

Participants were required to categorize circular square wave gratings into one of four categories based on their spatial frequency and orientation. The category boundaries were oblique to the two dimensions spanning the stimulus space, so that information from both dimensions had to be integrated predecisionally (Figure 5.1). This category structure was chosen as the acquisition of information-integration categories is known to be gradual and more dependent on training than the acquisition of easily verbalizable rules (Ashby & Maddox, 2005). Each category was composed of three subsets of stimuli which were distributed equidistantly across the stimulus space, resulting in several pairs of subsets of stimuli which had the same physical distance but were either on the same side of a category boundary or not. For each participant one of the four categories was randomly selected to be rewarded with a monetary gain for correct answers and punished with a monetary loss for wrong answers during training. This resulted in four equidistant pairs of stimulus subsets that crossed a category boundary that was reinforced, and four pairs of stimulus subsets that crossed a category boundary that was not reinforced. Using four categories instead of the usual two had the additional advantage that, while retaining a within-subject design, participants did not have to be retrained on other category boundaries within the same stimulus space. That is not only an unusual scenario in ecologically valid settings, but might also encourage the learning of flexible categorization rules dependent on the context of the task, rather than influencing stimulus representation.

To obtain stimuli, ordered pairs (x, y) were randomly sampled from the category distribution listed in Table 5.1. These pairs were subsequently transformed to determine the spatial frequency ($.83 + x/5.83$ cpd) and orientation ($0.35 + (y \times \pi/1500)$ radians) of a circular square-wave grating ($\sim 100\%$ contrast, 6° visual angle). Spatial phase was randomized across all stimuli to make local pixel intensity uninformative.

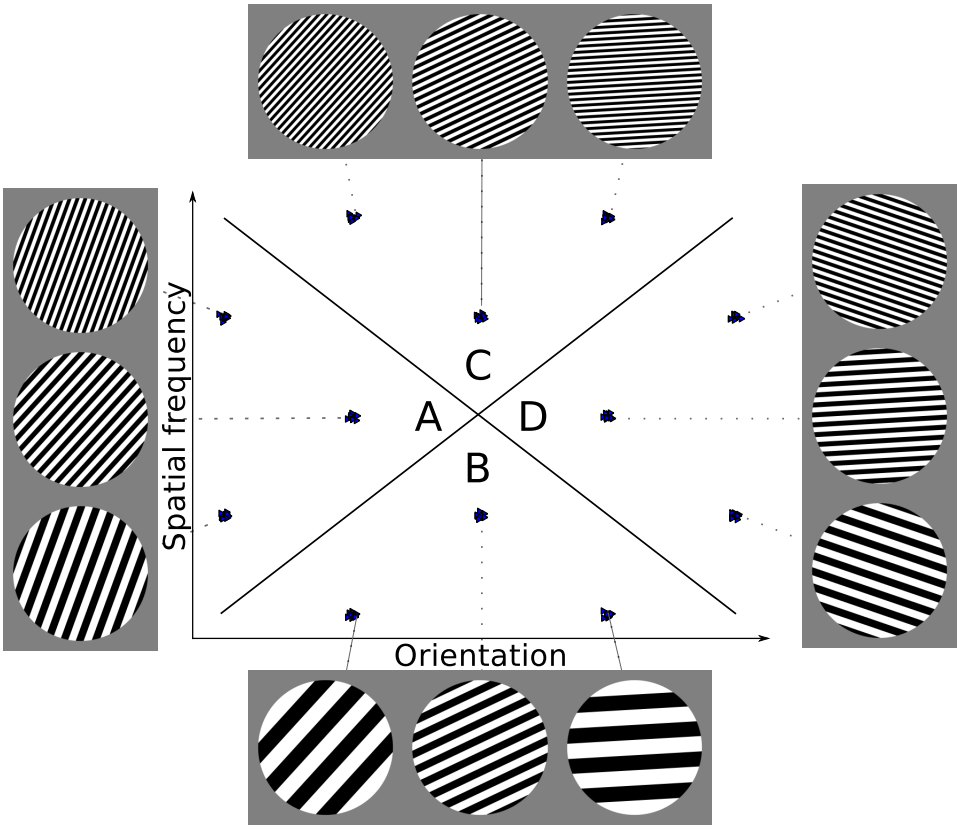

Figure 5.1. *Category structures and sample stimuli.* Each triangle denotes the orientation and spatial frequency of a single stimulus. The solid lines represent the category boundaries. Each category (A, B, C, and D) is composed of three subsets of stimuli. For all subsets an example stimulus is shown.

5.2.3. Training session

Procedure

Training was performed in a quiet, dimly lit room. All stimuli were presented on 50% gray background and with a circular 50% gray fixation dot of approximately 0.3° visual angle in diameter in the center. Before each trial a fixation cross was presented for 100 ms which was directly followed by the stimulus. The stimulus was presented for a maximum of 15 s or until the participant pressed one of four response keys to indicate category membership. After the response the stimulus was immediately removed and a feedback sound of 250 ms duration indicated the correctness of the response. The sound consisted of a high or low tone (900 or 350 Hz). For rewarded categories it was replaced by an equally long sound resembling a cash register for correct answers or a collision for wrong answers. The amount of earnings in cents was displayed below the stimulus in red letters. An adaptive training procedure based on models developed by Rescorla and Wagner (1972) and Gluck and Bower (1988b, 1988a) was used to ensure that all participants learned the complex category structure.

Table 5.1. *Distribution parameters*

	μ_x	μ_y	σ_x	σ_y	cov_{xy}
Category A					
set A1	5	190	5	5	0
set A2	5	560	5	5	0
set A3	190	375	5	5	0
Category B					
set B1	190	5	5	5	0
set B2	375	190	5	5	0
set B3	560	5	5	5	0
Category C					
set C1	190	745	5	5	0
set C2	375	560	5	5	0
set C3	560	745	5	5	0
Category D					
set D1	560	375	5	5	0
set D2	745	190	5	5	0
set D3	745	560	5	5	0

μ = mean for each dimension; σ = variance for each dimension; *cov* = covariance between dimensions.

Model for adaptive training

It was assumed that associative connections between members of each set of stimuli and the correct category response ($i = 1...12$ input nodes, $o = 1...4$ output nodes) are built up in a gradual manner according to an error correction rule:

$$\Delta V_{io} = \alpha R_o(\lambda_o - O_o)I_i$$

where V_{io} is the weight connecting the i-th input node with the o-th output node, α is the free learning rate parameter, R_o is the answer given by the participant, $\lambda = 0$ or 1 is the target value for the o-th output node as indicated by the feedback, and $I_i = 0$ or 1 represents activity on the i-th input node.

The activity O for each output node o on each trial was computed as:

$$O_o = \sum V_{io}I_o$$

with V_{io} being updated after each trial according to the error correction rule stated above.

Adaptive training

The free learning rate parameter was first set to a low level (0.03). After 120 trials this value was replaced with an estimate of the participant's learning rate based on the number of errors made during the first trials. To estimate the learning rate parameter the model was first run with all weights initialized to small random values (0-0.05) and run with different values for α. The model's response was determined to be the output node with the highest weight and the average number of errors over 500 instantiations of the model was taken as an estimate how many errors a participant with a certain learning rate would make. However, the learning rate was constrained to lie between 0.06 and 0.09 to make training duration comparable between participants, and no different learning rates were assigned to rewarded and not rewarded categories. The average learning rate determined by this procedure was 0.079 (SD = 0.015). After the setting of the individual learning rate parameter the experiment was divided into blocks of 100 trials and for each block the number of presented exemplars per category was determined in proportion to the weights of each input node to its corresponding output node. This procedure resulted in the presentation of more exemplars from input nodes that were weakly connected with their corresponding output nodes. After each block of 100 trials the training was paused until the participant indicated by a key press that he or she was ready to continue. Training ended when all correct weights were above 0.99 and was concluded by a final test block in which 10 exemplars from each of the twelve subsets of stimuli were presented.

Reinforcement

Correct answers to the rewarded category were rewarded while each failure to recognize exemplars of the rewarded category and each false alarm was punished. Reward and punishment were calculated in proportion to the association weights connecting each input node to the correct output node, so that in the beginning of the experiment reward amounted to 20 cents while punishment was zero. With increasing association strengths this relationship was gradually reversed. Maximum punishment was limited to 10 cents to shift the participants' focus to reward rather than punishment, while nevertheless preventing a bias for choosing the rewarded category. In the final test block all participants received a reward or punishment of 10 cents for each answer relating to the rewarded category.

5.2.4. FMRI session

Behavioral protocol

One day after the training session participants underwent an fMRI scanning session. Stimuli in the scanning session were identical to those in the training session but presented in a matrix of 3×3 gratings with a fixation dot only on the center stimulus. The presentation time for each stimulus was four seconds and interspersed with rest/fixation events of 104 s overall duration. The order of stimuli and null events was derived using the Optseq2 software (http://surfer.nmr.mgh.harvard.edu/optseq) and each run ended with 12 s of rest/fixation. To ensure continuous attention, participants were asked to press a button whenever one of the stimuli was different from the other eight. A high tone was presented for correct target detection and a low tone for missed targets. Within each of the seven runs five target stimuli were presented along with five homogenous displays from each subset. Target trials were excluded from further analysis.

FMRI Image acquisition and preprocessing

The parameters for fMRI image acquisition are detailed in Section 2.4. Seven runs of 190 functional images each were recorded. Processing of the functional images was performed using the FMRIB Software Library (FSL - Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Oxford, UK; http://www.fmrib.ox.ac.uk/fsl; Version 4.1), unless otherwise stated. Images were movement corrected to the middle image of the run and the estimated movement parameters were saved and later included in the statistical analysis. A temporal highpass filter of 1/128 Hz was applied to remove low frequency confounds. Images were spatially smoothed for the GLM analysis using a 5mm FWHM Gaussian kernel, while images later subjected to similarity analyses were not smoothed at this step. Twelve regressors representing the onsets and durations of stimulus presentation from each set of stimuli were fit to the data after convolution with a hemodynamic response function (gamma function with a standard deviation of 3 s and a mean lag of 6 s; temporal derivatives were added to the model).

Preparation of images for the similarity analysis

In the representational similarity analysis it is assumed that stimuli which are located at adjacent positions of the model space elicit BOLD signal patterns which are more similar to each other than signal patterns of stimuli which are more distant in the model space. To summarize the data for the similarity analysis, the contrast images for all twelve regressors from the first level analysis were subjected individually for each participant to a higher level fixed effects analysis. This resulted in maps representing the participant's mean activation over all seven runs for each subset of stimuli. The twelve maps with contrast parameter estimates resulting from this analysis were z-scored feature wise and submitted to the similarity analysis. To construct subsets of the imaging data for the calculation of distance matrices containing the correlational distances of each pair of stimuli, a searchlight approach (Kriegeskorte, Goebel, & Bandettini, 2006) was taken. A sphere-shaped neighborhood with 4 voxels radius ($n = 257$) was determined for each voxel within the whole brain using the Searchlight method implemented in PyMVPA (http://pymvpa.org; Hanke et al. (2009)). These empirical distance matrices were correlated with matrices containing the pairwise Euclidean distances between the stimuli in three different theoretical models of the stimulus space. The resulting coefficients were plotted back on the central voxel of the searchlight to build the correlation maps. All distance matrices were calculated using the hcluster package of SciPy (http://scipy.org; Jones et al. (2001)).

Model-based similarity analysis

Three models were developed to describe the similarity structure of the functional imaging data.

1. *Physical Model.* The Physical Model assumes that all distances between data points can be described by their Euclidean distances within the two-dimensional stimulus space.

2. *Categorical Model.* The Categorical Model also assumes that distances between data points represent Euclidean distances within the stimulus space. It, however, adds an additional free parameter: a scale factor which allows the members of all three subsets within a category to move closer to the centroid of the category. In this way distances within a category may decrease while distances between categories may increase.

3. *Reward Model.* The Reward Model is identical to the Category Model but adds a second free parameter. This allows the scale factor of the rewarded category to be different from the scale factor of the not rewarded categories. In this way, the representation of the subsets is allowed to be different within the rewarded category than in the other categories.

A visualization of all three models is provided in Figure 5.2. All free parameters were constrained to lie between 0 and 1 and the optimize package of SciPy was used to determine the parameters that maximize the correlation between each model and data ($df = 64$).

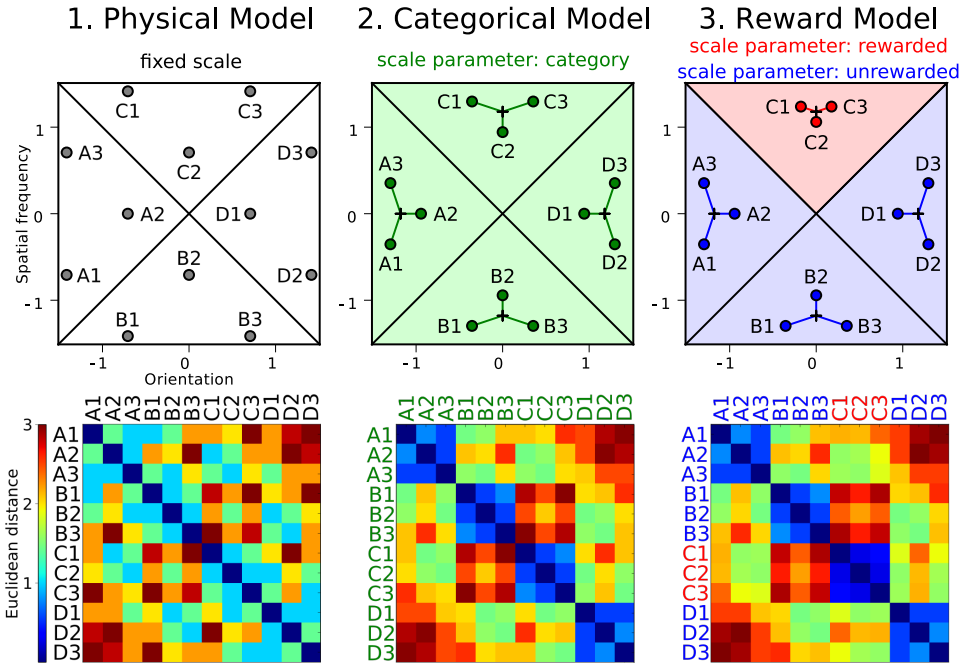

Figure 5.2. *Models of representational similarity.* Three different models were fit to the neuroimaging data. The Physical Model assumes that the similarity structure of the data is best described by the physical distances of each subset of stimuli (indicated by circles) within the two dimensional stimulus space. The Categorical Model introduces a free scale parameter that allows the similarity of stimuli within the same category to be higher than for stimuli in different categories. The scale factor in this example is 0.5, i.e. the distance between the category centroid (indicated by a cross) and the centroid of each subset of stimuli is scaled to be half the physical distance (indicated by green lines and background color). The Reward Model introduces an additional free scale parameter for the distances within the rewarded category (indicated by red lines and background color; blue lines and background color indicate the unrewarded categories). In the example shown the scale factor for the rewarded category is 0.25 (upper quadrant of the Reward Model), the scale factor for the unrewarded categories is 0.5. All models can be transformed to similarity matrices as shown in the second row of the figure, which were correlated with the empirical distance matrices obtained from the neuroimaging data.

Statistical testing of the similarity analysis

The Fisher's Z-transformed correlation maps of all three models with the neuroimaging data as well as the three maps containing the parameter estimates of the scale factors were linearly transformed to MNI space (Evans et al., 1993) and smoothed with a 5 mm FWHM Gaussian kernel. The correlation maps were tested using one-sample T-tests (5mm half width at half maximum (HWHM) variance smoothing) and all reported clusters survived a voxel-corrected threshold of $p < .005$ and an extent threshold of 50 contiguous voxels. As the models are nested, difference maps between them have a natural limit at zero. For small differences this might lead to non-normal distributions. Kolmogorov-Smirnov tests however confirmed that the hypothesis of normality cannot be rejected ($p < .2$, uncorrected) in the areas where significant differences were detected. Also, the introduction of additional free parameters will increase the correlation between model and data. Using our conservative threshold we determined the brain regions where this improvement was substantially above zero and thereby identified areas for further investigation of the degree of categorical representation. To this end the estimates of the scale factors in each model of each participant were extracted from the peak voxel of all significant clusters. Additionally, to specify the relationship to behavioral data from the training session, a median-split was used to group participants into high learners and low learners based on their performance in the final training block and included as between-subjects factor *learner group* into all analyses of the scale factor estimates.

Univariate GLM analysis

To check for univariate main effects of reward first level contrasts were set up comparing rewarded and unrewarded sets of stimuli individually for each participant and run. In an intermediate fixed effects analysis the contrast images from all seven runs were averaged for each participant and submitted to the higher level mixed effects analysis across participants.

5.3. Results

5.3.1. Behavioral Data

On average participants received 1210.3 (SD = 300.2) training trials, during which they earned an average of M = 21.15 € (SD = 4.94 €). In the final test block of 120 trials (10 trials from each subset) the average error rate was 9.43% (SD = 6.55), indicating that training was appropriate to ensure both a high level of performance and a sufficient number of error trials to detect a potential increase in signal detection for rewarded categories. With an average of .95 (SE = .19) the hit rate for rewarded categories was higher [T(15) = 3.06, $p < .05$] than the hit rate for not rewarded categories [M = .89; SE = .18], while the false alarm rate was lower [T(15) = -2.27; $p < .05$] for rewarded [M = .01, SE = .01] than unrewarded categories [M = .05, SE = .02]. Also, the average reaction time was lower [T(15) = -2.94, $p < .05$] for rewarded [M = 1084.66 ms, SE = 67.58 ms] than for not rewarded [M = 1325.08 ms, SE = 82.94 ms] categories. The behavioral data in the final test block therefore shows that despite otherwise identical training protocols, the categorization of items from the rewarded category was both more correct and faster than categorization of items from the not rewarded categories. During the fMRI session participants missed an average of 0.75 (SD = 1.4) target trials, indicating continuous attention to the stimuli.

5.3.2. FMRI session

Representational Similarity Analysis

1. Physical Model. Correlations of the similarity structure described by the Physical Model, which assumes that the distances between stimuli can be described by their Euclidean distances in the stimulus

space, and the correlation distance in the fMRI data were mainly observed throughout the occipital lobe with additional smaller clusters in the right anterior temporal fusiform cortex, the left posterior insular cortex and the left putamen (see Table 5.2 and Figure 5.3 A).

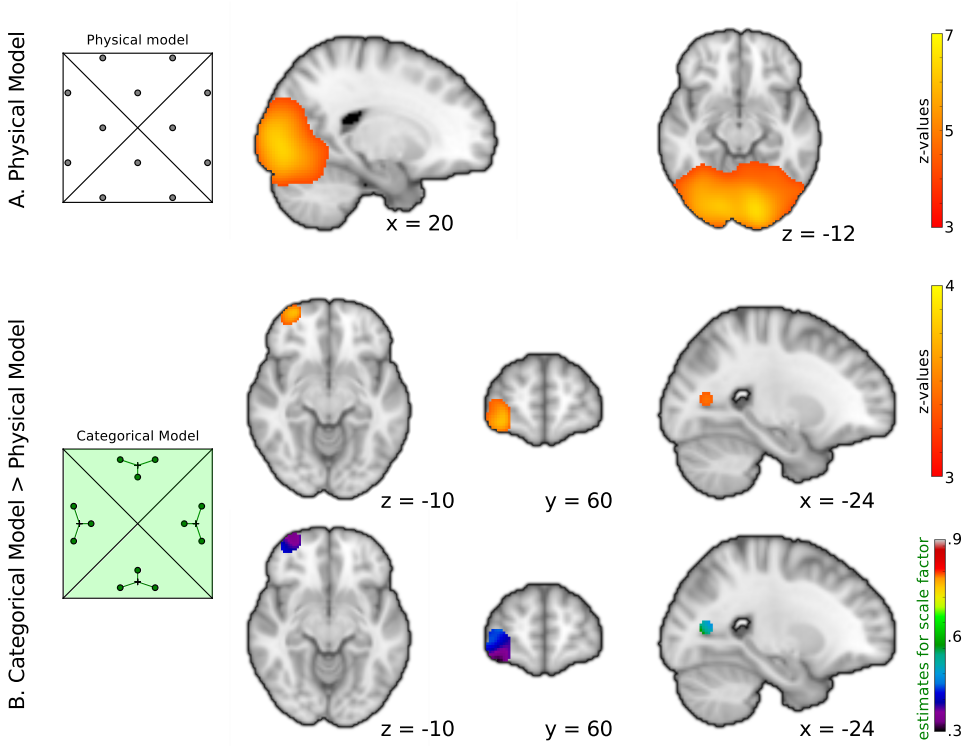

Figure 5.3. *Correlations of the Physical and Categorical Model.* **A.** Z-value map of the correlation of the similarity structure of the Physical Model with the similarity structure of the fMRI data. Pronounced correlations are observed throughout the occipital lobe. **B.** Z-value map showing where the Categorical Model correlated higher with the fMRI data than the Physical model. Two clusters in the occipital and the frontal cortex are observed. Additionally a map of the estimates of the scale factors in the Categorical Model, masked with the thresholded z-value map, is presented. Lower numbers indicate that the distance between the centroid of the category to each subset of stimuli was estimated to be shorter than the physical distance (see Figure 5.2). The left hemisphere is presented at the right side of coronal sections, and both z-value maps are thresholded at $p_{FWE} < .005$ (k >50).

2. Categorical Model > Physical Model. The Categorical Model adds a free parameter which allows the distances between members of the same category to be smaller than distances between members of different categories, irrespective of their physical distance. The addition of the free parameter led to significant increases in the correlation between model and data in the right frontal pole and the left rostral cuneus, bordering the calcarine sulcus and the retrosplenial cortex (see Table 5.2 and Figure 5.3 B). The scale factor estimates were extracted from the peak voxels of these two clusters and submitted to a two-factor ANOVA with repeated measures on the factor *cluster*, and *learner group* as between subjects factor. Although there was a trend indicating that the scale factor estimates were lower at the frontal pole than in the occipital cortex [F(1,14) = 2.91; p = .11], no significant effects were observed in the analysis of the scale factor estimates from the Categorical Model.

Table 5.2. *Correlation between models and data*

					MNI		
Region	Side	BA	k	Max. Z	x	y	z
1. Physical Model							
Occipital Cortex	17/18/19	b	25798	6.52	62	18	33
Anterior temporal fusiform cortex	20	r	437	3.81	32	64	16
Posterior insular cortex	41	l	311	3.64	-28	-36	10
Putamen		l	251	3.66	-22	0	4
2. Categorical Model > Physical Model							
Frontopolar cortex	10/11	r	650	3.74	32	60	-10
Inferior occipital cortex	17/18	l	135	3.45	-26	-70	10
3. Reward Model > Categorical Model							
Occipital Pole, extending to	17	b	6904	4.45	10	-96	-2
Lateral occipital cortex	19	r		4.17	42	-82	14
Lateral occipital cortex	19	l		4.01	-46	-84	-10
Cerebellum		b		4.17	-6	-54	-22
Superior occipital cortex	7	l	164	3.71	-14	-84	42
Superior occipital cortex	7	r	131	3.65	16	-74	52
Precentral Gyrus, extending to	6	r	785	3.95	24	-10	46
Postcentral Gyrus	4	r		3.67	16	-30	50
Postcentral Gyrus	4	l	641	4.04	-24	-22	32
Ventral Striatum		r	706	3.94	6	2	8
Temporal pole	20/38	r	370	3.72	34	16	-34
Amygdala/Hippocampus		l	332	3.63	-26	-12	-16
Posterior insular cortex	48	r	125	3.62	32	-30	26
Frontal orbital cortex	11	l	81	3.56	-26	40	8

l=left, r=right, b=bilateral, BA=Brodmann area, k=cluster size in voxels (2x2x2mm), max. Z=maximum Z-value

3. Reward Model > Categorical Model. The Reward Model adds a second free parameter which allows the scale factor of the rewarded category to be different from the scale factor of the not rewarded categories. Due to the introduction of this second parameter, significantly increased correlations between model and data were observed in occipital and parietal areas, the pre- and postcentral gyri (including the right Frontal Eye Field), the ventral striatum, the temporal pole, the amygdala, the posterior insular cortex and the frontal orbital cortex (see Table 5.2 and Figure 5.6). Please note that there was no overlap with the areas where activation patterns were best fit by the Categorical Model. The scale factor estimates from these clusters were submitted to a 10 (*cluster*) × 3 (*scale factor*) × 2 (*learner group*) ANOVA. A significant main effect of *cluster* [F(9, 126) = 2.13; $p < .05$] was observed. Further examination of the interaction between *cluster* and *scale factor* [F(11.08,155.11) = 2.09; $p < .05$] using post-hoc paired-samples T-tests showed that in early visual areas the scale factor estimate for the rewarded category was significantly lower than the estimate for the unrewarded category [T(15) = 2.25; $p < .05$], while in the orbitofrontal cortex this relationship was reversed [T(15) = -2.46; $p < .05$] (Figure 5.4). No other areas showed significant differences between scale factor estimates for rewarded and unrewarded categories in the post-hoc tests. Additionally, the mean scale factor estimates were lower [F(1,14) = 4.84; $p < .05$] for the group of high learners [M = .44, SE = .05] than for the group of low learners [M = .59, SE = .05] (Figure 5.5), indicating that in the group of high learners the stimuli in both rewarded and not rewarded categories were represented more similarly.

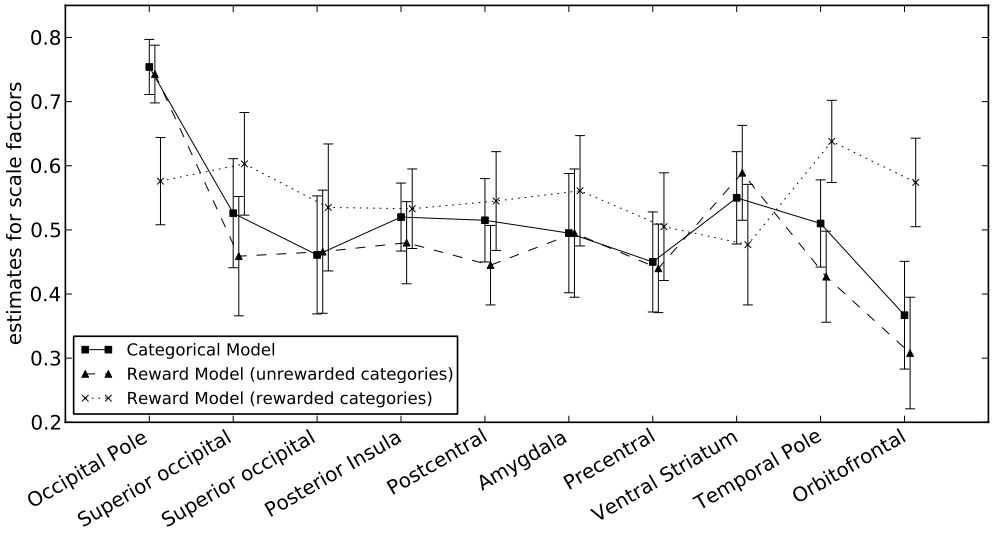

Figure 5.4. *Estimates of scale factors.* The parameter estimates represent the values of the scale factors that maximize the correlation between model and data in those areas where correlation was significantly higher with the Reward Model than the Categorical Model. Lower values indicate a higher within-category similarity. An interaction between *scale factor* and *cluster* is caused by lower estimates for the rewarded than unrewarded categories in occipital regions, and higher parameter estimates in the orbitofrontal cortex. No other clusters show significant differences between the parameter estimates for rewarded and unrewarded categories. Error bars represent the standard error of means.

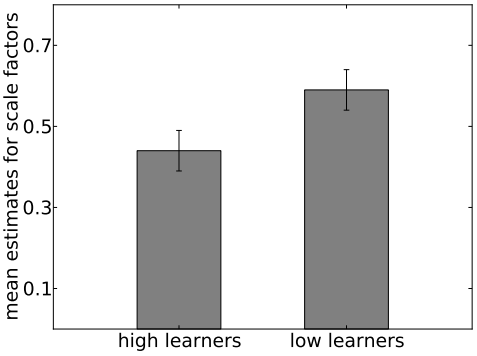

Figure 5.5. *Difference between high learners and low learners.* Participants were grouped into high learners and low learners using a median-split. The main effect of *learner group* in a 10(*cluster*)x3(*scale factor*)x2(*learner group*) ANOVA shows that in the group of high learners significantly lower overall estimates of the scale factor were observed, indicating that their representation of stimuli was more influenced by the category structure than in the group of low learners. No interaction with the cluster from which the estimates of the scale factors were extracted was observed, indicating that this effect was constant across all of the ten examined structures (compare Figure 5.4). Error bars represent the standard error of means.

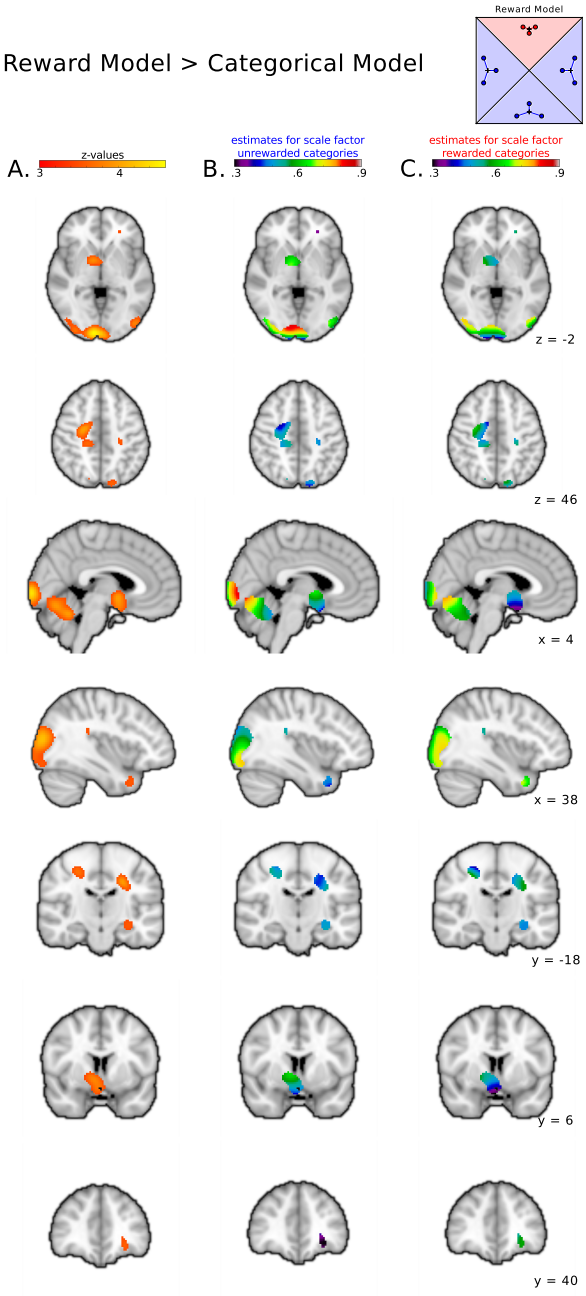

Figure 5.6. *Correlations of the Reward Model.* **A.** Z-Value map showing where the Reward Model correlated higher with the FMRI data than the Categorical Model, thresholded at $p_{FWE} < .005 (k > 50)$. The Reward Model has two free parameters, one for the rewarded category and one for the unrewarded categories. **B.** Map of the estimates of the scale factor in the Reward Model for unrewarded categories, masked with the thresholded z-value map from A. **C.** Map of the estimates of the scale factor in the Reward Model for the rewarded category, masked with the thresholded z-value map from A. Lower numbers in the scale factor estimate maps indicate that the distance between the centroid of the category to each subset of stimuli was estimated to be shorter than the physical distance (see Figure 5.2). The left hemisphere is presented at the right side of coronal sections.

Univariate Results

In a standard GLM analysis comparing rewarded stimuli with unrewarded stimuli no univariate effect was observed at a threshold of $p_{corrected} < .05$, indicating that in the absence of a behavioral task and with no probability of reward the formerly rewarded sets of stimuli did not evoke a response that differed in height from the response to unrewarded sets of stimuli within one voxel.

5.4. Discussion

To investigate the lasting influence of categorization training on representation a categorization task with equidistant subsets of stimuli was developed. The subsets were either divided by a category boundary or not, and responses with regard to two of the four boundaries were monetarily reinforced during training. In a wide network including visual and reward-related areas the representation of stimuli after categorization training reflected both their category membership and the reward associated with them. Irrespective of the physical distance of the stimuli, and in the absence of a categorization task, the representation of each subset of stimuli was more similar for members of the same category than for members of different categories. Both for monetarily reinforced and cognitively trained categories this effect was stronger in participants who had learned the task better. Additionally, in early visual areas the within-category similarity of stimulus representation was higher for the category whose boundaries were monetarily reinforced, while in the OFC the relationship was reversed. This suggests that the higher incentive salience induced by monetary reward as compared to cognitive feedback leads to decreased perceptual noise at the category boundaries in sensory processing areas. In contrast, members of the rewarded category are processed more distinctly in the OFC, an area known to encode subjective value representations and complex information necessary to define what possible outcomes are predicted by the state the environment is in (Wallis, 2007; O'Doherty, 2007).

5.4.1. Areas correlated with the theoretical models

To obtain these findings the similarity matrix predicted by three theoretically derived models of the stimulus space was correlated with the similarity matrix obtained from fMRI data (Kriegeskorte, Mur, & Bandettini, 2008). The first model specified the actual physical distances in the two-dimensional stimulus space. Significant correlations with this model were mainly observed throughout the occipital cortex. It is well-known that the physical characteristics of visual input are processed in occipital areas (Brewer, Liu, & Wade, 2005; Holmes, 1945). Therefore this finding supports the validity and sensitivity of the method employed.

A parameter allowing for both lower within-category and higher between-category distances was estimated in the second model. This model described the fMRI activations better than the physical model in the rostral part of the cuneus, bordering retrosplenial cortex, which is possibly a human homologue of the area prostriata (Rosa, Casagrande, Preuss, & Kaas, 1997; Sanides, 1972), and a larger area in the lateral frontopolar cortex (lFPC). The lFPC is well-known to support episodic memory retrieval (Gilbert et al., 2006), while the retrosplenial cortex and adjacent parts of the medial occipital cortex are functionally connected with the hippocampus (Kahn, Andrews-Hanna, Vincent, Snyder, & Buckner, 2008). They have been shown to be involved in both episodic memory and spatial navigation tasks (Ranganath, Heller, Cohen, Brozinsky, & Rissman, 2005; Vann, Aggleton, & Maguire, 2009; von Cramon, 1992; Wolbers & Büchel, 2005). In the present experiment participants were not required to categorize the stimuli during fMRI data acquisition, nor was there another task with high attentional demands. It is, however, possible that participants were either consciously or automatically remembering the learned classifications. In any case, our findings establish that category structure is represented in two areas that support memory even in the absence of an explicit categorization task.

Subsequently, a third model was fit to the fMRI data which did not only take into account the category structure of the presented stimuli, but also the reinforcement associated with them. It allowed the

distances in the rewarded category to be different from the distances in the unrewarded categories, and fit the fMRI data best in a widespread network. According to their functional relevance the observed areas can be aggregated into three major groups: areas involved in visual processing (occipital pole region, striate and peristriate cortex, superior occipital cortex), in motor responses (pre- and postcentral gyrus), and in the processing of reward value associated with stimuli (ventral striatum, orbitofrontal cortex, amygdala/hippocampus). All areas of this network have previously been implicated in category learning (see Section 1.2).

The result that category membership in combination with reward influences stimulus representations mainly in early visual areas complements extensive previous research on learning-related representational changes along the ventral visual processing stream. Category specific neuronal response patterns are often observed in more ventral visual areas along the inferior temporal cortex (Freedman, 2008; Sigala & Logothetis, 2002). However, training related changes in the visual system are highly dependent on the complexity of the stimuli and the duration of training, with higher visual areas in the fusiform gyrus being implicated in the processing of highly complex stimuli after extensive training, while earlier visual areas are implicated in the processing of simple visual stimuli and after less extensive training (Freedman, 2008; Gauthier, Wong, & Palmeri, 2010; Kourtzi et al., 2005; Li et al., 2007; Sigala, Gabbiani, & Logothetis, 2002). Presently, effects of reward and category membership were strongest in early visual areas, which is in agreement with our simple visual stimuli. They were defined by a single orientation and spatial frequency, which both are processed already in V1 (Mazer, Vinje, McDermott, Schiller, & Gallant, 2002).

One important difference between animal studies and studies in humans is that animals generally receive explicit reinforcement, while humans are mostly trained based on cognitive feedback. Sigala et al. (2002) directly compared the effects of categorization training using these two different procedures. Their monkeys received juice rewards and showed more pronounced behavioral effects of category training, i.e. increased within-category similarity ratings, than the human participants, who received only correctness feedback. This is in line with our findings in that within the same human participants a stronger effect of category for rewarded stimuli was observed both in the behavioral data and in functional activation in visual areas.

Activation patterns in several areas known to be implicated in reward processing were also best explained by a model accounting for both the category structure and the reinforcement structure associated with the stimuli. The ventral striatum plays a central role in the processing of salient stimuli (Zink et al., 2004) as well as in the in expectation of rewards and the processing of outcomes that are better than expected (Niv, 2009). In the present experiment rewards were only associated with the stimuli during training, but not during data acquisition, indicating that information about reward associated with the stimuli is also present in the ventral striatum when currently no reward can be obtained. A patch of highest correlation with the reward model was also observed in the OFC, an area known to represent information about the value of expected rewards based on the current state (O'Doherty, 2007; Knutson, Delgado, & Phillips, 2008), as well as in the left amygdala and hippocampus. Both areas are known to play a role in aversive conditioning (Büchel & Dolan, 2000; Sehlmeyer et al., 2009); specifically the hippocampus has been suggested to underlie associative category learning and the acquisition of equivalence between stimuli (Coutureau et al., 2002; Preston, Shrager, Dudukovic, & Gabrieli, 2004).

5.4.2. Examination of the parameter estimates

Despite the fact that the presented network representing both category and reward structure fits in well with previous observations, the identification of a network of brain areas where correlation between models and data is significantly increased by extending the models has to be interpreted with care in the absence of an untrained control group. However, this approach is well suited to identify areas of interest for further examination. Importantly, the method employed in the present study yields interpretable estimates of within-category similarity in the patterns of functional activation. Comparing the estimates

for reinforced and not reinforced categories in areas of interest reveals that in early visual processing areas within-category similarity is higher in the reinforced categories. Previous research has established that stimulus-reward pairings influence response properties of neurons in early sensory areas (Bao, Chan, & Merzenich, 2001; Frankó, Seitz, & Vogels, 2010). The present findings complement these results by showing that the similarity structure of stimulus representation after categorization training is categorical, and that this is true especially for reinforced category boundaries. An amplified representation of feature values close to category boundaries induced by perceptual learning is functionally advantageous in categorization tasks, as it supports different behavioral reactions in response to small changes in sensory input (Roelfsema et al., 2010). That this effect is stronger for the rewarded category fits in well with the higher functional relevance of correct categorizations for these stimuli.

In contrast to the pattern in early sensory areas, in the OFC activation patterns associated with members of the reinforced category are less similar than activation patterns associated with non-reinforced stimuli. This indicates a qualitatively special processing of stimuli that have previously been associated with reward. The OFC is known to support the disambiguation of similar environmental states by representing complex information about the subjective value of stimuli and the potential outcomes associated with them. It is involved in behavioral choice based on both present internal states and knowledge of the structure of the decision problem, and has been suggested to supply a working memory for value representations that allows to anticipate the consequences of behavior (O'Doherty, 2007; Takahashi et al., 2011; Wallis, 2007). The present results expand this knowledge by establishing that more valuable stimuli are processed distinctly within the OFC. While items that were not reinforced during training are represented categorically, stimuli that have previously been associated with reward are processed individually. This processing advantage allows a preferential examination of potentially valuable stimuli and therefore enhances the representation of the decision space to account for its reward structure. Previous accounts of categorical representation have assumed that training, i.e. the repeated exposure to stimuli and their association with reward, shapes the neuronal representation in sensory areas (Sigala & Logothetis, 2002; Kim & Bao, 2008). Only in the presence of a categorization task, these signals are read out by category selective neurons in the lateral prefrontal cortex (Freedman, 2008; Moore et al., 2006; Jiang et al., 2007). In accordance with these findings, we observed a categorical pattern of representation in sensory areas and no category information in lateral prefrontal areas during passive viewing. However, the OFC as only prefrontal area did represent categorical information in the absence of a categorization task. Importantly, this information included the representation of information on potential value in a way that can be utilized to maximize the outcome of decisions.

5.4.3. Potential sources of the effects of reward

To discuss the potential underlying causes of the influence of reward observed in the present study it is important to note that they cannot be explained by physical differences between the stimuli, as the rewarded category was randomly chosen for each subject. However, several other partially overlapping general accounts have to be considered when interpreting the findings. First, monetarily rewarding correct responses leads to better task performance. In the present experiment, error rates at the end of training were 6% lower for the rewarded category than for the unrewarded category. Although this is a weak behavioral effect, it might account for the better fit of the reward model in areas with increased within-category similarity in the rewarded category, i.e. in visual areas, as indicated by the lower scale factor estimates of participants who had learned the task better. This account implies that the effects of reward observed in the present study are mediated by a faster learning of reinforced stimulus-response associations and should be alleviated with further training when task performance is perfect across all categories. Second, the rewarded category is more salient, and therefore stimuli from it may have attracted more attentional resources. Effects of attention and reward are often confounded and seem to share several neuronal substrates, including orbitofrontal regions and the ventral striatum (Maunsell, 2004; Stoppel et al., 2011; Zink et al., 2004). It however has to be noted that the presently observed pattern of areas best fit by a reward model does not correspond to the parietal and prefrontal

areas which are traditionally assumed to form the attention network (Corbetta & Shulman, 2002; Mesulam, 1999; Posner & Petersen, 1990). Third, reward may specifically affect neuronal representations. Neuromodulatory systems, which broadly project to most cortical areas, change their activity in response to reward prediction and outcome. Although other neuromodulatory systems are known to be involved in plasticity (Aston-Jones & Cohen, 2005; Bakin & Weinberger, 1996; Richardson & DeLong, 1986), the role of dopamine in learning is particularly well established (Niv, 2009). In accordance with this account, the pairing of a stimulus with transient dopaminergic signals has been shown to change the representation of that stimulus in early sensory areas (Bao et al., 2001). As the presented study was designed to provide an overview of all areas that might be influenced by the general effects of reward, the specific source of the effect has to be addressed using more specialized experimental paradigms.

5.4.4. Summary

During passive viewing, the similarity pattern of fMRI responses was observed to be best fit by a model accounting for both the category and the reward structure associated with the stimuli in a widespread network. This network includes sensory, (pre)motor and reward-related areas, all of which have previously been reported to support category learning (see Section 1.2). An examination of the parameter estimates of the model reveals that this effect is stronger in participants who had learned the task better, and that in visual areas members of the rewarded category are represented more similarly, while in the OFC rewarded stimuli are processed more distinctly. The results of the present experiment thereby provide evidence for a fine-tuning of representation in sensory areas to ensure high task performance, while the OFC specifically represents environmental states leading to differently valued possible future outcomes.

6 Summary and General Conclusions

The experiments presented in this thesis examined the impact of different forms of outcome information on the neural substrates of category learning. A review of the literature on category learning shows that the presence, timing and nature of outcome information play a central role especially for complex category learning tasks, which require the integration of information over several trials and a predecisional integration of information within stimuli (Ashby & Maddox, 2005). Optimal performance in these tasks, which rely on a gradual acquisition of stimulus-response-outcome associations, critically depends on subcortical structures, mainly the striatum (Shohamy et al., 2008). In studies on category learning outcome information has been operationalized as information about the correctness of a decision, which is referred to as cognitive feedback in this thesis. The effects of cognitive feedback have been suggested to be mediated by the mesencephalic dopaminergic system, which has previously mainly been examined in the context of reward-based learning.

A wealth of electrophysiological animal research on the neural substrates of reward-learning has established the central role of the mesencephalic dopaminergic system. Dopaminergic midbrain neurons have been shown to encode the expectation of rewards and deviations from this expectation, the reward prediction error (Schultz, 2007). A parallel line of research in humans has examined this system using fMRI. It has to be emphasized that fMRI is mainly sensitive to local field potentials which reflect integrated activity over large neural populations (Logothetis, 2003) and is not sensitive to the neurotransmitter system conveying the signals. However, due its importance for the investigation of learning, the neurochemistry underlying fMRI activations in dopaminergic midbrain nuclei and the ventral striatum during reward learning has been explicitly addressed. Results indicate that fMRI activity in these areas is likely related to dopaminergic activity during reward learning (Knutson & Gibbs, 2007; Pessiglione et al., 2006; Schott et al., 2008). FMRI studies on reward-based learning in humans, which often focus on gambling paradigms with monetary rewards, have shown similar patterns of activation in dopaminergic target areas as predicted by electrophysiological results (O'Doherty et al., 2007). Dopaminergic target areas have also been shown to be involved in tasks where no explicit rewards are provided (Aron et al., 2004). In ecologically valid settings many decision are not followed by reward or even external information on correctness (Hogarth, 2006) and the role of dopaminergic target structures under these conditions is not well specified.

The set of studies presented in this thesis compared the influence of three different forms of outcome information, (1) monetary reward, (2) cognitive information, and (3) internal signals on correctness, on the neural substrates of category learning and representation of learned information as measured with fMRI in healthy young participants.

6.1. Summary of the experimental procedures and results

In the first experiment the effects of monetary reward and cognitive feedback were directly compared within the same participants. To this end, two parallel information-integration category tasks were developed, which the participants performed in a single event-related fMRI session. Monetary rewards were provided for one of the task versions, in other version information about the correctness of the response was provided. Similar functional activations in dopaminergic target areas were observed in both task versions, including activation in the putamen for successful as compared to unsuccessful categorization, and in the nucleus accumbens and the body of the caudate nucleus for positive as compared to negative feedback. In contrast, a single structure in the ventral striatum, the nucleus accumbens, showed higher activation during monetary reward anticipation compared to the anticipation of cognitive feedback. This activation was influenced by the motivational state of the participants. It was predicted by measures of intrinsic motivation in the cognitive feedback task and by measures of extrinsic motivation in the rewarded task.

The results of this first experiment confirm that the processing of monetary reward and cognitive feedback share similar neural substrates. However, in many learning environments information on probable outcome has to be internally constructed (Henriksson et al., 2010). The neural correlates of internally generated signals on correctness in the complete absence of feedback were examined in a second experiment using the same task. During test trials after observational training a direct comparison between correct and wrong responses reveals that activation in dopaminergic target areas, including the nucleus accumbens, putamen and dmPFC, is modulated by the correctness of the answer similarly as during feedback-based training. As in reinforcement learning fMRI activation in dopaminergic target areas is observed in response to errors in reward prediction rather than to the absolute value of reward (Schultz, 2007), a standard reinforcement learning model was introduced in which subjective confidence ratings acquired after each trial served as outcome measure. The fMRI activation was correlated with this prediction error in the striatum (nucleus accumbens and putamen). Activation in dopaminergic projection sites therefore follows the same pattern in response to prediction errors on confidence as it does do during reinforcement learning in response to reward prediction errors, but with respect to internally generated signals based on knowledge of the structure of the environment. Furthermore, ventral striatal activation was higher for novel stimuli, which might support the allocation of attention to unfamiliar stimuli.

In a final step, the lasting impact of the type of reinforcement during training on the representation of categories was addressed. Participants were trained to categorize stimuli into one of four categories, while only answers with regard to one category were monetarily rewarded. For the other categories information on correctness was provided. The similarity structure of fMRI data acquired after successful training during passive viewing was best fit by a model accounting for both the category membership and reward structure of stimuli in sensory, (pre)motor and reward-related areas. The representation of category exemplars was more similar for those participants who had learned the categorization task best and was more similar the rewarded category in early visual areas. In contrast, rewarded stimuli were processed more distinctly in the OFC, another prominent target area of the dopaminergic system.

6.2. Contributions and implications

Within the last decade, broad-ranging efforts have been made to extend results on reward-based learning and the role of dopamine therein from animal studies to the human domain. Research on humans using fMRI could replicate many of the findings from animal studies. The striatum has been shown to respond to reward value and the prediction error both in classical and instrumental conditioning tasks (O'Doherty et al., 2007). The studies presented in this thesis help to transfer these results to more ecologically valid settings, in which the category membership and therefore functional relevance of stimuli is not always clear, and after many decisions reward or outcome information is not available.

The presented results support the assumption that forms of learning which depend on response contingent feedback rely on related neuronal substrates as reward learning (Ashby & Maddox, 2005; Nomura & Reber, 2008; Seger, 2008) by showing that dopaminergic target areas are activated in a very similar way during both forms of training. However, this finding has to be differentiated for different subcomponents of the striatum. In Experiment 1 we provide evidence that a single structure, the nucleus accumbens, responds with pronouncedly higher activation during the expectation of monetary reward as compared to the expectation of cognitive feedback. Previous findings on monetary rewards suggest that the nucleus accumbens codes for the expected positive incentive properties of a reward (Cooper, 2008; Knutson et al., 2001). Our observations expand these findings by showing that the accumbens also responds differentially when comparing qualitatively different types of outcome information, i.e. cognitive feedback and monetary reward. The ventral striatum as part of the motivational corticostriatal loop has been shown to exert directed influence on the visual cortex during categorization, and it has been suggested to participate in learning by providing the motivational background for decisions (Lopez-Paniagua & Seger, 2011). This fits in well with the present results showing that the pattern of activation within the nucleus accumbens is influenced by the motivational state of the participant. Potentially due to the utilization of two tasks that were trained in parallel, no differences in performance in response to the type of outcome information were observed. However, reward is known to influence behavior during learning in other paradigms (see Experiment 3). The present results provide evidence for differential activation patterns within the nucleus accumbens depending on the type of outcome information, which might play a role in mediating these behavioral differences.

These findings are extended by the second experiment which demonstrates possible mechanisms by which learning in the absence of external rewards can be supported by the striatum. In the complete absence of feedback, striatal activation follows a similar pattern in responding to the prediction error as it does during reinforcement learning, but with respect to signals based on subjective confidence, i.e. activation is higher whenever confidence is higher than expected. A potential source for these activations lies within the error monitoring system (Ridderinkhof, Ullsperger, et al., 2004). Internally detected errors might be interpreted either as source of information on performance outcome similarly as errors detected due to external feedback, or indicate that the state of the environment, which signals the probability of future reward, has changed. We show that the process leads to an overall higher activation in striatal nuclei in response to correct compared to incorrect task performance which parallels the activation pattern observed in response to external performance feedback. Additionally, when no actions are performed and stimuli are passively observed, ventral striatal activation decreases with stimulus novelty. Both processes might support the acquisition of category structure in the absence of feedback. These results are especially interesting as patients with Parkinson's disease, which is characterized by a loss of dopaminergic input mainly to the dorsal striatum (Schönberg et al., 2010), are impaired on feedback-based complex category learning tasks, while observational learning is spared (Shohamy et al., 2004). This has been interpreted to show that observational learning is possible without the recruitment of striatal structures (Shohamy et al., 2008), which is in contrast to our results indicating that healthy young participants do recruit the striatum in at least two distinct ways, namely in response to novelty and internal signals on confidence about the decision. The latter process can support learning but, in the absence of external information, also strengthen false associations. Whether each of the processes is necessary or sufficient for optimal task performance and how the processing of internal and external outcome information interact in ecologically valid settings, where outcome information is partially available, is subject to further research.

To examine lasting influences of the type of outcome information during training on the representation of stimuli, we combined theoretical modeling based on previous behavioral observations (Goldstone, 1994; Sigala et al., 2002) with an analysis based on the multivariate similarity structure of fMRI activation patterns (Kriegeskorte, Mur, & Bandettini, 2008). We identified the network of brain areas in which representation is influenced both by category membership and reward, and observed differential effects within these areas. The network representing categorical information includes early visual and reward-related subcortical and prefrontal areas, and is therefore more widespread than previously assumed

(Freedman, 2008; Li et al., 2009). Moreover, by showing that better learned rewarded categories are represented more similarly in visual areas, our finding supports previous accounts suggesting that perceptual learning is involved in categorical perception (Notman et al., 2005; Roelfsema et al., 2010), and that visual areas encode mainly perceived form similarity (Freedman, 2008; Li et al., 2007). A major novel finding is that within the OFC rewarded stimuli are processed more distinctly, indicating a preferential processing of potentially valuable environmental stimuli within this area. As indicated by higher estimates of within-category similarity and between-category separation observed in participants who have learned the task better, these distributed and differential neural representations can support perceptual and cognitive processes leading to a categorical representation of the environment based on stimulus-reward contingencies, and thereby can be instrumental in maximizing the outcome of behavioral decisions.

6.3. Conclusions

The experiments presented in this thesis show that reward-based learning and learning in the absence of explicit rewards share common basic neural substrates. Specifically, major dopaminergic target areas including the striatum respond to the expectation of cognitive feedback similarly as to the expectation of monetary reward and, in the complete absence of external outcome information, activation in the striatum follows a prediction error on confidence. Further similarities are observable even after training, when categorical structure, which has been acquired during reward-based and feedback-based learning, is concurrently represented in visual areas. However, both quantitative and qualitative differences can be observed between reward-based learning, cognitive feedback-based learning and learning in the absence of reward. During category learning effects in the striatum are smaller in the absence of explicit rewards, and in visual areas the representation of the categorical structure of stimuli is less pronounced after feedback-based than after reward-based training. Also, the motivational states induced by the type of outcome information influence the observed pattern of activations, and a specific representation of potentially more valuable stimuli in orbitofrontal areas can be observed. By showing analogies to reward-based learning the present results provide a parsimonious account of category learning in the absence of explicit feedback, however also shed light on the neural substrates of differential behavioral effects observed due to differing training protocols.

References

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *NeuroImage*, *31*(2), 790–795.

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–517.

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, *32*(3), 537–351.

Aizenstein, H. J., Macdonald, A. W., Stenger, V. A., Nebes, R. D., Larson, J. K., Ursu, S., et al. (2000). Complementary Category Learning Systems Identified Using Event-Related Functional MRI. *Journal of Cognitive Neuroscience*, *12*(6), 977–987.

Allwood, C. M., Granhag, P. A., & Johansson, H. (2000). Realism in confidence judgements of performance based on implicit learning. *European Journal of Cognitive Psychology*, *12*(2), 165–188.

Aron, A. R., Shohamy, D., Clark, J., Myers, C. E., Gluck, M. A., & Poldrack, R. A. (2004). Human Midbrain Sensitivity to Cognitive Feedback and Uncertainty During Classification Learning. *Journal of Neurophysiology*, *92*(2), 1144–1152.

Arthurs, O. J., & Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*, *25*(1), 27–31.

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449–483). Hillsdale, NJ: Lawrence Erlbaum.

Ashby, F. G. (2011). *Statistical Analysis of fMRI Data*. Cambridge, MA: MIT Press.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481.

Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 1–36). New York, NY: Elsevier.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 33–53.

Ashby, F. G., & Maddox, W. T. (1998). Stimulus Categorization. In M. H. Birnbaum (Ed.), *Measurement, judgment, and decision making, Handbook of perception and cognition* (2nd ed., pp. 251–301). San Diego, CA: Academic Press.

Ashby, F. G., & Maddox, W. T. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition*, *30*(5), 666–677.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149–178.

Ashby, F. G., & O'Brien, J. B. (2007). The effects of positive versus negative feedback on information-integration category learning. *Attention, Perception, & Psychophysics*, *69*(6), 865.

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011). COVIS. In E. M. Pothos & Wills A. J. (Eds.), *Formal approaches in categorization* (pp. 65–87). New York, NY: Cambrigde University Press.

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, *61*(6), 1178–1199.

Ashby, F. G., & Valentin, V. V. (2005). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 548–574). New York, NY: Elsevier.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.

Bakin, J. S., & Weinberger, N. M. (1996). Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(20), 11219–11224.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience*, *27*(31), 8161–8165.

Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. (1992). Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, *25*(2), 390–397.

Bao, S., Chan, V. T., & Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature*, *412*(6842), 79–83.

Baranski, J. V., & Petrusic, W. M. (1999). Realism of confidence in sensory discrimination. *Perception & Psychophysics*, *61*(7), 1369–1383.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391–431.

Bertsekas, D. P., & Tsitsiklis, J. N. (1996). *Neuro-dynamic Programming*. London, UK: Athena.

Bjork, J. M., & Hommer, D. (2007). Anticipating instrumentally obtained and passively-received rewards: a factorial fMRI investigation. *Behavioural brain research*, *177*(1), 165–170.

Boettiger, C. a., & D'Esposito, M. (2005). Frontal networks for learning and executing arbitrary stimulus-response associations. *The Journal of Neuroscience*, *25*(10), 2723–2732.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*(2), 619–639.

Brewer, A. A., Liu, J., & Wade, A. R. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, *8*(8), 1102–1209.

Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., & Boussaoud, D. (2008). Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory. *Cerebral Cortex*, *18*(7), 1485–1495.

Büchel, C., & Dolan, R. J. (2000). Classical fear conditioning in functional neuroimaging. *Current Opinion in Neurobiology*, *10*(2), 219–223.

Bunzeck, N., & Düzel, E. (2006). Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron*, *51*(3), 369–379.

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*, 69–84.

Buxton, R. B., Uludağ, K., Dubowitz, D. J., & Liu, T. T. (2004). Modeling the hemodynamic response to brain activation. *NeuroImage*, *23*(Suppl 1), S220–233.

Cincotta, C. M., & Seger, C. A. (2007). Dissociation between striatal regions while learning to categorize via feedback and via observation. *Journal of Cognitive Neuroscience*, *19*(2), 249–265.

Cohen, H., & Lefebvre, C. (Eds.). (2005). *Handbook of Categorization in Cognitive Science*. Oxford, UK: Elsevier.

Cohen, M. X. (2008). Neurocomputational mechanisms of reinforcement-guided learning in humans: A review. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(2), 113–125.

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., & Schlaepfer, T. E. (2009).

Neuroelectric signatures of reward learning and decision-making in the human nucleus accumbens. *Neuropsychopharmacology*, *34*(7), 1649–1658.

Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *The Journal of Neuroscience*, *27*(2), 371–378.

Cooper, J. C. (2008). Valence and salience contribute to nucleus accumbens activation. *NeuroImage*, *39*(1), 538–547.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215.

Coutureau, E., Killcross, A. S., Good, M., Marshall, V. J., Ward-Robinson, J., & Honey, R. C. (2002). Acquired equivalence and distinctiveness of cues: II. Neural manipulations and their implications. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 388–396.

Craig, A. D. B. (2009). How do you feel–now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), 59–70.

Daniel, R. (2008). *Networks involved in the processing of different levels of uncertainty during prototype-distortion learning.* Unpublished dipoma thesis, Friedrich Schiller Univerisity, Jena.

Daniel, R., & Pollmann, S. (2010). Comparing the neural basis of monetary reward and cognitive feedback during information-integration category learning. *The Journal of Neuroscience*, *30*(1), 47–55.

Daniel, R., & Pollmann, S. (2012). Striatal activations signal prediction errors on confidence in the absence of external feedback. *NeuroImage*, *59*(4), 3457–3467.

Daniel, R., Wagner, G., Koch, K., Reichenbach, J. R., Sauer, H., & Schlösser, R. G. M. (2011). Assessing the neural basis of uncertainty in perceptual category learning through varying levels of distortion. *Journal of Cognitive Neuroscience*, *23*(7), 1781–1793.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319*(5867), 1264–1267.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In *Decision Making, Affect, and Learning: Attention and Performance XXIII.* Oxford, UK: Oxford University Press.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

Daw, N. D., & Shohamy, D. (2008). The Cognitive Neuroscience of Motivation and Learning. *Social Cognition*, *26*(5), 593–620.

Dayan, P., & Balleine, B. W. (2002). Reward, Motivation and Reinforcement Learning. *Neuron*, *36*(2), 285–298.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current Opinion in Neurobiology*, *18*(2), 185–196.

de Brujin, E. R. A., de Lange, F. P., von Cramon, D. Y., & Ullsperger, M. (2009). When errors are rewarding. *The Journal of Neuroscience*, *29*(39), 12183–12186.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *The Journal of Neuroscience*, *25*(50), 11730–11737.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, *62*(1), 119–142.

DeGutis, J., & D'Esposito, M. (2007). Distinct mechanisms in visual category learning. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(3), 251–259.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, *8*(11), 1611–1618.

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to

reward and punishment: effects of valence and magnitude manipulations. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(1), 27–38.

Delgado, M. R., Miller, M. M., Inati, S., & Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *NeuroImage*, *24*(3), 862–873.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the Hemodynamic Responses to Reward and Punishment in the Striatum. *Journal of Neurophysiology*, *84*(6), 3072–3077.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of Neuroscience*, *4*(8), 2051–2062.

D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Experimental Brain Research*, *133*(1), 3–11.

Dreher, J.-C., Kohn, P., & Berman, K. F. (2006). Neural coding of distinct statistical properties of reward information in humans. *Cerebral Cortex*, *16*(4), 561–573.

Düzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B. C., Schott, B. H., & Tobler, P. N. (2009). Functional imaging of the human dopaminergic midbrain. *Trends in Neurosciences*, *32*(6), 321–328.

Eger, E., Ashburner, J., Haynes, J.-D., Dolan, R. J., & Rees, G. (2008). fMRI activity patterns in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, *20*(2), 356–370.

Elliott, R., Newman, J. L., Longe, O. A., & William Deakin, J. F. (2004). Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *NeuroImage*, *21*(3), 984–990.

Elwin, E., Juslin, P., Olsson, H., & Enkvist, T. (2007). Constructivist coding: learning from selective feedback. *Psychological Science*, *18*(2), 105–110.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *Nuclear Science Symposium and Medical Imaging Conference, IEEE Conference Record*, *3*(31), 1813–1817.

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001). A possible role of the striatum in linear and nonlinear category learning: Evidence from patients with Huntington's disease. *Behavioral Neuroscience*, *115*(4), 786–798.

Frahm, J., Bruhn, H., Merboldt, K. D., & Hänicke, W. (1992). Dynamic MR imaging of human brain oxygenation during rest and photic stimulation. *Journal of Magnetic Resonance Imaging*, *2*(5), 501–505.

Franke, G. (1995). *Die Symptom-Checkliste von Derogatis - Deutsche Version*. Göttingen, Germany: Beltz Test Gesellschaft.

Frankó, E., Seitz, A. R., & Vogels, R. (2010). Dissociable Neural Effects of Long-term Stimulus–Reward Pairing in Macaque Visual Cortex. *Journal of Cognitive Neuroscience*, *22*(7), 1425–1439.

Freedman, D. J. (2008). Neural mechanisms of visual categorization: insights from neurophysiology. *Neuroscience & Biobehavioral Reviews*, *32*(2), 311–329.

Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature*, *443*(7107), 85–88.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A Comparison of Primate Prefrontal and Inferior Temporal Cortices during Visual Categorization. *Journal of Neuroscience*, *23*(12), 5235–5246.

Gauthier, I., Wong, A. C.-N., & Palmeri, T. J. (2010). Manipulating visual experience: comment on Op de Beeck and Baker. *Trends in Cognitive Sciences*, *14*(1), 22–30.

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self-and externally caused failure of goal achievement. *NeuroImage*, *47*(4), 2023–2030.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., et al. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *Journal of Cognitive Neuroscience*, *18*(6), 932–948.

Gillebert, C. R., Op de Beeck, H. P., Panis, S., & Wagemans, J. (2009). Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences. *Journal of Cognitive Neuroscience*, *21*(6), 1054–1064.

Gläscher, J., Daw, N. D., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595.

Gluck, M. A., & Bower, G. H. (1988a). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, *27*(2), 166–195.

Gluck, M. A., & Bower, G. H. (1988b). From conditioning to category learning: an adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–247.

Goense, J. B. M., & Logothetis, N. K. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology*, *18*(9), 631–640.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, *5*(1), 10–16.

Gold, J. I., & Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *The Journal of Neuroscience*, *23*(2), 632–651.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*(2), 178–200.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Gottfried, J. A., O'Doherty, J. P., & Dolan, R. J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, *22*(24), 10829–10837.

Gottfried, J. A., O'Doherty, J. P., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*, *301*(5636), 1104–1107.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron*, *49*(5), 757–763.

Haber, S. N., & Knutson, B. (2009). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4–26.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*(2), 148–154.

Han, S., Huettel, S. A., Raposo, A., Adcock, R. A., & Dobbins, I. G. (2010). Functional significance of striatal responses during episodic decisions: recovery or goal attainment? *The Journal of Neuroscience*, *30*(13), 4767–4775.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*(1), 37–53.

Hare, T. A., O'Doherty, J. P., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience*, *28*(22), 5623–5630.

Harnad, S. (2005). To cognize is to categorize: Cognition is categorization. *Handbook of categorization in cognitive science*, 20–45.

Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of Neurophysiology*, *95*(2), 948–959.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., et al. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *The Journal of Neuroscience*, *24*(7), 1660–1665.

Heldmann, M., Rüsseler, J., & Münte, T. F. (2008). Internal and external information in error processing. *BMC Neuroscience*, *9*(33).

Helie, S., Roeder, J. L., & Ashby, F. G. (2010). Evidence for cortical automaticity in rule-based categorization. *The Journal of Neuroscience*, *30*(42), 14225–14234.

Henriksson, M. P., Elwin, E., & Juslin, P. (2010). What is coded into memory in the absence of outcome feedback? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 1–16.

Hogarth, R. M. (2006). Is confidence in decisions related to feedback? Evidence from random samples of real-world behavior. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 456–484). New York, NY: Cambridge University Press.

Holmes, G. (1945). The organization of the visual cortex in man. *Proceedings of the Royal Society of London*, *132*(869), 348–361.

Holroyd, C. B., & Coles, M. G. H. (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error Related Negativity. *Psychological Review*, *109*(4), 679–709.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L. E., Mars, R. B., Coles, M. G. H., et al. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*(5), 497–498.

Holroyd, C. B., Yeung, N., Coles, M. G. H., & Cohen, J. D. (2005). A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology: General*, *134*(2), 163–191.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*. Sunderland, MA: Sinauer Associates.

James, W. (1890). *The principles of psychology*. New York: Holt.

Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., et al. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Human Brain Mapping*, *28*(4), 294–302.

Jezzard, P., Matthews, P. M., & Smith, S. M. (Eds.). (2001). *Functional MRI: An introduction to methods*. Oxford, UK: Oxford University Press.

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., Vanmeter, J., & Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron*, *53*(6), 891–903.

Jocham, G., & Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neuroscience & Biobehavioral Reviews*, *33*(1), 48–60.

Joel, D., & Wiener, I. (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience*, *96*(3), 451–474.

Jones, E., Oliphant, T., & Peterson, P. (2001). *SciPy: Open source scientific tools for Python.* http://www.scipy.org/.

Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., & Buckner, R. L. (2008). Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*(1), 129–139.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York, NY: Appleton-Century-Crofts.

Kanwisher, N. G., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.

Kayser, C., Kim, M., Ugurbil, K., Kim, D.-S., & König, P. (2004). A comparison of hemodynamic and neural responses in cat visual cortex using complex stimuli. *Cerebral Cortex*, *14*(8), 881–891.

Kim, H., & Bao, S. (2008). Distributed representation of perceptual categories in the auditory cortex. *Journal of Computational Neuroscience*, *24*(3), 277–290.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, *308*, 78–83.

Knutson, B., Delgado, M. R., & Phillips, P. E. M. (2008). Representation of subjective value in the striatum. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 389–406). New York, NY: Academic Press.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*(17), 3683–3687.

Knutson, B., & Gibbs, S. E. B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology*, *191*(3), 813–822.

Knutson, B., Taylor, J., & Kaufman, M. (2005). Distributed neural representation of expected value. *The Journal of Neuroscience*, *25*(19), 4806–4812.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*(5648), 1181–1185.

Kourtzi, Z., Betts, L. R., Sarkheil, P., & Welchman, A. E. (2005). Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biology*, *3*(7), e204.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(4), 1–28.

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*(6), 1126–1141.

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, *72*(5), 341–72.

Kwong, K. K. (1992). Dynamic Magnetic Resonance Imaging of Human Brain Activity During Primary Sensory Stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(12), 5675–5679.

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, *23*(4), 727–738.

Li, S., Mayhew, S. D., & Kourtzi, Z. (2009). Learning shapes the representation of behavioral choice in the human brain. *Neuron*, *62*(3), 441–452.

Li, S., Mayhew, S. D., & Kourtzi, Z. (2011). Learning Shapes Spatiotemporal Brain Patterns for Flexible Categorical Decisions. *Cerebral Cortex*, doi:10.1093/cercor/bhr309.

Li, S., Ostwald, D., Giese, M., & Kourtzi, Z. (2007). Flexible coding for categorical decisions in the human brain. *The Journal of Neuroscience*, *27*(45), 12321–12330.

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex*, *15*(10), 1621–1631.

Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, *7*, 274–281.

Little, D. M., & Thulborn, K. R. (2005). Correlations of cortical activation and behavior during the application of newly learned categories. *Brain Research*, *25*(1), 33–47.

Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 732–753.

Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *The Journal of Neuroscience*, *23*(10), 3963–3971.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.

Logothetis, N. K. (2010). Neurovascular Uncoupling: Much Ado about Nothing. *Frontiers in Neuroenergetics*, *2*(June), 1–4.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, *66*, 735–69.

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(22), 9493–9498.

Lopez-Paniagua, D., & Seger, C. a. (2011). Interactions within and between corticostriatal loops during component processes of category learning. *Journal of Cognitive Neuroscience*, *23*(10), 3068–3083.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309–332.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*(1), 49–70.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 650–662.

Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 100–107.

Maddox, W. T., Love, B. C., & Glass, B. D. (2008). When more is less: Feedback effects in perceptual category learning. *Cognition*, *108*(2), 578–589.

Mahon, B. Z. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annual Review of Psychology*, *60*, 27–51.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.

Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences*, *8*(6), 261–5.

Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H., & Gallant, J. L. (2002). Spatial frequency and orientation tuning dynamics in area V1. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(3), 1645–1650.

McAuley, E. (1989). Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339–346.

Meeter, M., Radics, G., Myers, C. E., Gluck, M. A., & Hopkins, R. O. (2008). Probabilistic categorization: how do normal participants and amnesic patients do it? *Neuroscience & Biobehavioral Reviews*, *32*(2), 237–248.

Menon, M., Jensen, J., Vitcu, I., Graff-Guerrero, A., Crawley, A. P., Smith, M. A., et al. (2007). Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biological Psychiatry*, *62*(7), 765–772.

Mesulam, M. M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *354*(1387), 1325–1346.

Metereau, E., & Dreher, J.-C. (2012). Cerebral Correlates of Salient Prediction Error for Different Rewards and Punishments. *Cerebral Cortex*, doi: 10.1093/cercor/bhs037.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Mizuno, K., Tanaka, M., Ishii, A., Tanabe, H. C., Onoe, H., Sadato, N., et al. (2008). The neural basis of academic achievement motivation. *NeuroImage*, *42*(1), 369–378.

Montague, P. R., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Comparative Neurology*, *16*(5), 1936–1947.

Moore, C. D., Cohen, M. X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *The Journal of Neuroscience*, *26*(43), 11187–11196.

Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., & Malach, R. (2005). Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. *Science*, *309*(5736), 951–954.

Münte, T. F., Heldmann, M., Hinrichs, H., Marco-Pallares, J., Krämer, U. M., Sturm, V., et al. (2007). Nucleus Accumbens is Involved in Human Action Monitoring: Evidence from Invasive Electrophysiological Recordings. *Frontiers in Human Neuroscience*, *1*(11).

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(49), 20911–20916.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *The Journal of Neuroscience*, *32*(2), 551–562.

Niv, Y., & Montague, P. R. (2008). Theoretical and Empirical Studies of Learning. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 329–350). New York, NY: Academic Press.

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., et al. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex*, *17*(1), 37–43.

Nomura, E. M., & Reber, P. J. (2008). A review of medial temporal lobe and caudate contributions to visual category learning. *Neuroscience & Biobehavioral Reviews*, *32*(2), 279–291.

Notman, L. A., Sowden, P. T., & Ozgen, E. (2005). The nature of learned categorical perception effects: a psychophysical approach. *Cognition*, *95*(2), B1–14.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, *14*, 769–776.

O'Doherty, J. P. (2007). Lights, camembert, action! The role of human orbitofrontal cortex in encoding stimuli, rewards, and choices. *Annals of the New York Academy of Sciences*, *1121*, 254–72.

O'Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*, *49*(1), 157–166.

O'Doherty, J. P., Dayan, P., Friston, K. J., Critchley, H. D., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337.

O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, *33*(5), 815–826.

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, *1104*, 35–53.

Ogawa, S. (1992). Intrinsic Signal Changes Accompanying Sensory Stimulation: Functional Brain Mapping with Magnetic Resonance Imaging. *Proceedings of the National Academy of Sciences*, *89*(13), 5951–5955.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*(6), 419–427.

Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *The Journal of Neuroscience*, *26*(50), 13025–13036.

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*(2), 97–98.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045.

Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Moyano, J. C., Myers, C. E., et al. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of Functional MRI Data Analysis*. New York, NY: Cambrigde University Press.

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., et al. (2005). The neural correlates of motor skill automaticity. *The Journal of Neuroscience*, *25*(22), 5356–5364.

Poldrack, R. A., & Willingham, D. T. (2006). Functional neuroimaging of skill learning. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (Second ed., pp. 113–148). Cambridge, MA: The MIT Press.

Pollmann, S., & Maertens, M. (2005). Shift of activity from attention to motor-related brain areas during visual learning. *Nature Neuroscience*, *8*(11), 1494–1496.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, *14*(2), 148–152.

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*(3), 381–390.

Raichle, M. E., & Mintun, M. A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience*, *29*, 449–476.

Raiguel, S., Vogels, R., Mysore, S. G., & Orban, G. A. (2006). Learning to see the difference specifically alters the most informative V4 neurons. *The Journal of Neuroscience*, *26*(24), 6589–6602.

Ranganath, C., Heller, A., Cohen, M. X., Brozinsky, C. J., & Rissman, J. (2005). Functional connectivity with the hippocampus during successful memory formation. *Hippocampus*, *15*(8), 997–1005.

Rangel, A., Camerer, C. F., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998a). Contrasting cortical activity associated with category memory and recognition memory. *Learning & Memory*, *5*(6), 420–428.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998b). Cortical areas supporting category learning identified using functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(2), 747–750.

Redgrave, P., Gurney, K., & Reynolds, J. (2008). What is reinforced by phasic dopamine signals? *Brain Research Reviews*, *58*(2), 322–339.

Rescorla, R. A. (1970). Reduction in the effectiveness of reinforcement after prior excitatory conditioning. *Learning and Motivation*, *1*(4), 372–381.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II, Current Research and Theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Richardson, R. T., & DeLong, M. R. (1986). Nucleus basalis of Meynert neuronal activity during a delayed response task in monkey. *Brain Research*, *399*(2), 364–368.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The Role of the medial frontal cortex in cognitive control. *Science*, *306*, 443–447.

Ridderinkhof, K. R., Wildenberg, W. P. M. van den, Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefontal cortex in action selection, response inhibition, performance monitoring, and reward based learning. *Brain and Cognition*, *56*(2), 129–140.

Rodriguez, P. F., Aron, A. R., & Poldrack, R. A. (2006). Ventral–striatal/nucleus–accumbens sensitivity to prediction errors during classification learning. *Human Brain Mapping*, *27*(4), 306–313.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. a. (2001). Factors that determine false recall: a multiple regression analysis. *Psychonomic Bulletin & Review*, *8*(3), 385–407.

Roelfsema, P. R., Ooyen, A. van, & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences*, *14*(2), 64–71.

Romo, R., Hernández, A., & Zainos, A. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron*, *41*(1), 165–173.

Rosa, M. G. P., Casagrande, V. A., Preuss, T., & Kaas, J. H. (1997). Visual Field Representation in Striate and Prestriate Cortices of a Prosimian Primate (Galago garnetti). *Journal of Neurophysiology*, *77*(6), 3193–3217.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397.

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. J. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*(6), 1054–1069.

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *The Journal of Neuroscience*, *30*(40), 13525–13536.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, *43*(3), 450–461.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67.

Sanides, F. (1972). Representation in the cerebral cortex and its areal lamination patterns. *Structure and function of nervous tissue*, *5*, 329–453.

Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *The Journal of Neuroscience*, *27*(47), 12860–12867.

Schönberg, T., O'Doherty, J. P., Joel, D., Inzelberg, R., Segev, Y., & Daw, N. D. (2010). Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson's disease patients: evidence from a model-based fMRI study. *NeuroImage*, *49*(1), 772–781.

Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., et al. (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *The Journal of Neuroscience*, *28*(52), 14311–14319.

Schott, B. H., Niehaus, L., Wittmann, B. C., Schütze, H., Seidenbecher, C. I., Heinze, H. J., et al. (2007). Ageing and early-stage Parkinson's disease affect separable neural mechanisms of mesolimbic reward processing. *Brain*, *130*(Pt9), 2412–2424.

Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology*, *80*, 1–27.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*(3), 199–207.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115.

Schultz, W. (2007). Behavioral dopamine signals. *Trends in Neurosciences*, *30*(5), 203–210.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions*, *6*(24).

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Seger, C. A. (2006). The basal ganglia in human learning. *Neuroscientist*, *12*(4), 285–290.

Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience & Biobehavioral Reviews*, *32*(2), 265–278.

Seger, C. A., & Cincotta, C. M. (2005). The Roles of the Caudate Nucleus in Human Classification Learning. *The Journal of Neuroscience*, *25*(11), 2941–2951.

Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual Review of Neuroscience*, *33*, 203–219.

Seger, C. A., Peterson, E. J., & Cincotta, C. M. (2010). Dissociating the Contributions of Independent Corticostriatal Systems to Visual Categorization Learning Through the Use of Reinforcement Learning Modeling and Granger Causality Modeling. *NeuroImage*, *50*(2), 644–656.

Seger, C. A., Poldrack, R. A., Prabhakaran, V., Zhao, M., Glover, G. H., & Gabrieli, J. D. E. (2000). Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia*, *38*(9), 1316–1324.

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., et al. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS One*, *4*(6), e5865.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, *60*(6), 1169–1181.

Sescousse, G., Redouté, J., & Dreher, J.-C. (2010). The architecture of reward value coding in the human orbitofrontal cortex. *The Journal of Neuroscience*, *30*(39), 13095–13104.

Seymour, B., Daw, N. D., Dayan, P., Singer, T., & Dolan, R. J. (2007). Differential encoding of losses and gains in the human striatum. *The Journal of Neuroscience*, *27*(18), 4826–4831.

Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., et al. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, *429*(10), 664–667.

Shen, L., & Alexander, G. E. (1997). Neural correlates of a spatial sensory-to-motor transformation in primary motor cortex. *Journal of Neurophysiology*, *77*(3), 1171–1194.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain*, *127*(Pt 4), 851–859.

Shohamy, D., Myers, C. E., Kalanithi, J., & Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience & Biobehavioral Reviews*, *32*(2), 219–236.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*(2), 378–389.

Sigala, N., Gabbiani, F., & Logothetis, N. K. (2002). Visual categorization and object representation in monkeys and humans. *Journal of Cognitive Neuroscience*, *14*(2), 187–198.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, *415*(6869), 318–320.

Smith, B. W., Mitchell, D. G. V., Hardin, M. G., Jazbec, S., Fridberg, D., Blair, R. J. R., et al. (2009). Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *NeuroImage*, *44*(2), 600–609.

Smith, E. E. (2008). The case for implicit category learning. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 3–16.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*(3), 315–327.

Stoppel, C. M., Boehler, C. N., Strumpf, H., Heinze, H. J., Hopf, J.-M., & Schoenfeld, M. A. (2011). Neural processing of reward magnitude under varying attentional demands. *Brain Research*, *1383*, 218–229.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*(1), 9–44.

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 497–537). Cambridge, MA: MIT Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Takahashi, Y. K., Roesch, M. R., Wilson, R. C., Toreson, K., O'Donnell, P., Niv, Y., et al. (2011).

Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neuroscience*, *14*(12), 1590–1597.

Tobler, P. N., Dickinson, A., & Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *The Journal of Neuroscience*, *23*(32), 10402–10410.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, *97*(2), 1621–1632.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, *41*(2), 281–292.

Ullsperger, M. (2006). Performance monitoring in neurological and psychiatric patients. *International Journal of Psychophysiology*, *59*(1), 59–69.

Ullsperger, M. (2010). Genetic association studies of performance monitoring and learning from feedback: the role of dopamine and serotonin. *Neuroscience & Biobehavioral Reviews*, *34*(5), 649–659.

Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *The Journal of Neuroscience*, *27*(15), 4019–4026.

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, *10*(11), 792–802.

Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain. *Neuron*, *72*(1), 166–177.

Vitay, J., Fix, J., Beuth, F., Schroll, H., & Hamker, F. H. (2009). Biological models of reinforcement learning. *Künstliche Intelligenz*, *3*, 12–18.

Vogels, R. (1999). Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *European Journal of Neuroscience*, *11*(4), 1239–1255.

Vogels, R., Sary, G., Dupont, P., & Orban, G. A. (2002). Human brain regions involved in visual categorization. *NeuroImage*, *16*(2), 401–414.

von Cramon, D. Y. (1992). The septo-hippocampal pathways and their relevance to human memory: A case report. *Cortex*, *28*(3), 411–422.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43–48.

Waldschmidt, J. G., & Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *NeuroImage*, *56*(3), 1791–802.

Wallis, J. D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annual Review of Neuroscience*, *39*(1), 538–547.

Watkins, C. (1989). *Learning from delayed rewards*. Unpublished doctoral dissertation, Cambridge University, Cambridge.

Westphal, S. (2009). *Observationales Lernen visueller Kategorien mit und ohne motorische Assoziation*. Unpublished Bachelor's thesis, Otto-von-Guericke University, Magdeburg.

White, C. M., & Koehler, D. J. (2004). Missing information in multiple-cue probability learning. *Memory & Cognition*, *32*(6), 1007–1018.

Wise, R. A. (1989). The brain and reward. In J. M. Liebman & S. J. Cooper (Eds.), *The neuropharmacological basis of reward* (pp. 377–424). Oxford, UK: Oxford University Press.

Wittmann, B. C., Bunzeck, N., Dolan, R. J., & Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage*, *38*(1), 194–202.

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Düzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*(3), 459–467.

Wolbers, T., & Büchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *The Journal of Neuroscience*, *25*(13), 3333–3340.

Worsley, K., Marrett, S., & Neelin, P. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73.

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable

systems for gain- and loss-related value predictions and errors of prediction in the human brain. *The Journal of Neuroscience*, *26*(37), 9530–9537.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience*, *24*(28), 6258–6264.

Zakay, D., & Tuvia, R. (1998). Choice latency times as determinants of post-decisional confidence. *Acta Psychologica*, *98*(1), 103–115.

Zhang, J., Riehle, A., Requin, J., & Kornblum, S. (1997). Dynamics of Single Neuron Activity in Monkey Primary Motor Cortex Related to Sensorimotor Transformation. *Cortex*, *17*(6), 2227–2246.

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, *42*(3), 509–517.

A Experiment 1: Supplemental results

A.1. Experiment 1: Results of the pilot study

Details on the method of the pilot study for Experiment 1 are available in Section 3.2.

A.1.1. First session: Training duration

An ANOVA with repeated measures on *task* (first / second task) with the between-subjects factors *order of reward* (first task rewarded / not rewarded), *order of stimuli* (first task circles / lines) and *order of category structure* (slope of first task positive / negative) was calculated to compare the number of trials participants needed to reach criterion (80% correct trials within a single block of 50 trials) in each task. Four participants had to be excluded from this analysis as they did not reach criterion within the maximum training duration of five blocks. No significant main effects of the between-subjects factors or interactions with the within-subjects factor were observed on the .2 level, indicating that all four task versions are parallel in terms of learning speed. Results of the ANOVA are detailed in Table A.1.

Table A.1. *Experiment 1: ANOVA results for the pilot study (first session)*

Source	F(1,7)	p
task	0.83	.39
order of reward	0.39	.55
order of stimuli	0.39	.55
order of category structure	0.58	.47
task × order of reward	0.08	.79
task × order of stimuli	0.07	.80
task × order of category structure	0.65	.45

A.1.2. Second session: Error rates

An ANOVA with repeated measures on *task* (first / second task) with the between-subjects factors *order of reward* (first task rewarded / not rewarded), *order of stimuli* (first task circles / lines) and *order of category structure* (slope of first task positive / negative) was calculated to compare the average error rates in the first and second task in the second session of the pilot study to Experiment 1. No significant

main effects of the between-subjects factors or interactions with the within-subjects factor were observed on the .2 level, indicating that all task versions are parallel in terms of error rates after training. Results of the ANOVA are detailed in Table A.2.

Table A.2. *Experiment 1: ANOVA results for the pilot study (second session)*

Source	F(1,10)	p
task	0.05	.83
order of reward	0.06	.81
order of stimuli	0.35	.57
order of category structure	0.06	.81
task × order of reward	0.68	.43
task × order of stimuli	0.26	.62
task × order of category structure	1.66	.23

A.2. Experiment 1: Results of the model-based analysis of the behavioral data

This section reports supplemental results of the model fitting procedures that were applied to each participant's behavioral data in Experiment 1 (see Chapter 3). Details on the method are provided in Section 2.3 and for illustration examples for the best fitting models of four single subjects are given in Figure A.1. The resulting BIC values are summarized in Table A.3.

Table A.3. *Experiment 1: BIC values*

	Monetary reward					*Cognitive feedback*				
Subject	II[a]	L[b]	O[c]	C1[d]	C2[e]	II	L	O	C1	C2
1	262	397	549	398	**217**	222	451	382	**162**	390
2	414	504	727	405	**388**	341	302	728	371	**261**
3	**258**	399	526	308	339	255	394	561	**202**	371
4	313	469	596	**290**	469	**283**	398	621	486	327
5	419	484	649	653	**352**	**324**	407	683	350	449
6	209	425	520	509	**132**	234	480	521	**134**	381
7	**252**	346	534	314	401	277	388	559	**192**	384
8	**170**	431	447	186	415	332	339	647	371	**245**
9	295	520	701	415	**252**	201	407	432	**195**	427
10	**397**	446	489	427	598	**436**	450	619	488	565
11	380	460	659	355	490	295	392	500	478	**275**
12	364	476	655	333	501	**359**	505	399	425	513
13	**284**	441	600	377	311	**184**	363	724	235	442
14	250	261	632	**213**	248	302	681	490	**254**	502
15	184	399	370	**152**	393	265	448	398	524	**196**
16	**299**	479	394	403	329	402	348	635	419	**215**

[a] II = information-integration model
[b] L = unidimensional model based on linewidth
[c] O = unidimensional model based on orientation
[d] C1 = first conjunctive model
[e] C2 = second conjunctive model

Smaller values indicate better model fit. The smallest values for each subject and task are highlighted.

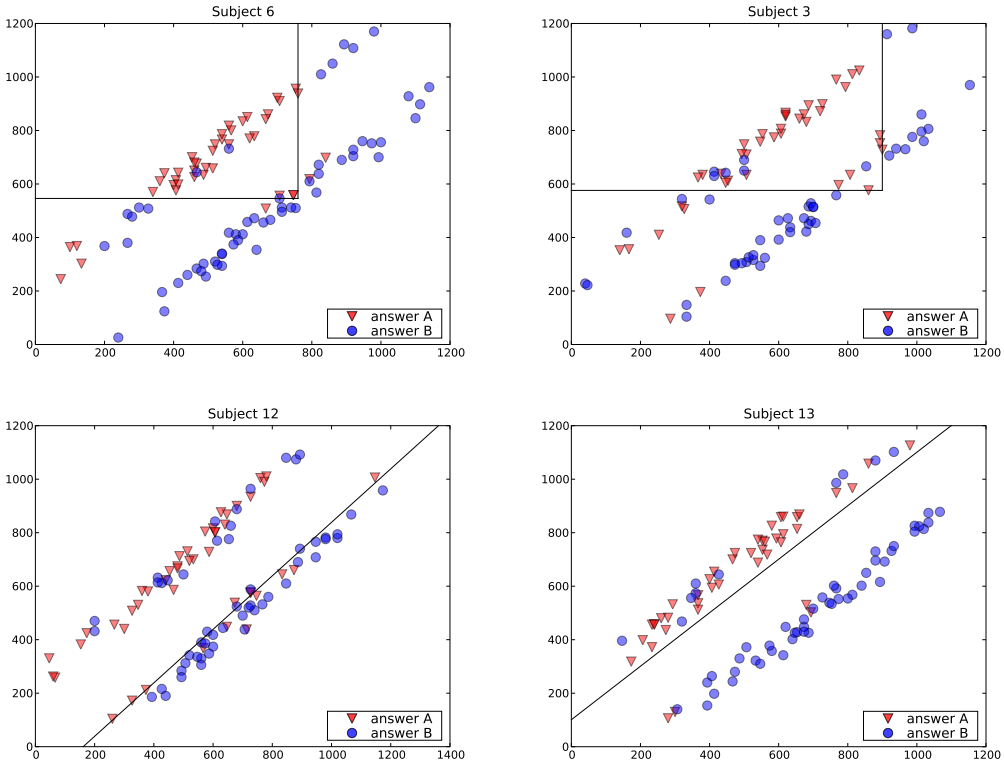

Figure A.1. *Best fitting models for four subjects.* To illustrate the results of the model fitting procedures, for both the conjunction models and the information-integration model those two subjects were selected that show the biggest difference in BIC values between best fitting and second best fitting model for the task with a positive slope of the decision bound. The values along the axes represent x and y values sampled from the bivariate normal distributions specified in Table 3.1. Stimuli which were categorized to be from category "A" are marked by red triangles, stimuli which were categorized to be from category "B" are marked by blue squares. The optimal decision bound for the category structure is depicted in Figure 3.1. The solid black lines indicate the subjective decision bound of the participant as estimated by the best fitting model.

B Experiment 2: Supplemental methods and results

B.1. Experiment 2: Pilot study

This section reports the main results of the pilot study to Experiment 2, which was performed by Westphal (2009). Twenty-five participants with an average age of 24.1 years [range = 20-31, SD = 3.2; one excluded due to misunderstanding the instructions] (see also Section 2.1) participated in the pilot study for Experiment 2. Stimuli are described in Section 4.2.2, and only behavioral data was acquired. The task was presented in four blocks, each consisting of 100 trials of observation and 100 trials of testing. During observational trials half of the participants responded to the stimuli by pressing the response key associated with the presented category label, while the other half of the participants always pressed the same key independent of the label. In the final testblock participants reached an accuracy of M = 77.0 % (SD = 16.5). An ANOVA with repeated measures on *block* and *response condition* as between-subjects factor showed a significant training effect [$F(3,66) = 11.3$, $p < .001$], but no effect of *response condition* [$F(1,22) = 0.3$, $p = .85$] or a *block* $\times$ *response condition* interaction [$F(3,66) = 0.0003$, $p = .99$] (Figure B.1). The results of this pilot study reveal that behavioral improvements are observable when training is performed in the absence of feedback, however after 400 observational training trials performance is still as low as 77%. Additionally, associating the stimuli with a response key during observation does not influence the training effect. Therefore in the main study reported in Chapter 4 further training sessions were introduced, while participants were not required to press response keys during observational trials.

B.2. Experiment 2: Results of the model-based analysis of the behavioral data

This section reports supplemental results on the model fitting procedures that were applied to each participant's behavioral data in Experiment 2 (see Chapter 4). Details on the method are provided in Section 2.3. The resulting BIC values are summarized in Table B.1.

Table B.1. *Experiment 2: BIC values (continued on next page)*

	First session (fMRI)					*Second session (training)*				
Subject	II[a]	L[b]	O[c]	C1[d]	C2[e]	II	L	O	C1	C2
1	299	532	747	**290**	596	**30**	432	711	212	351
2	463	546	658	**425**	523	**120**	499	727	299	302
3	330	396	439	**313**	413	**123**	302	752	238	276
4	517	565	**511**	533	561	551	460	453	**326**	534
5	428	263	514	**274**	366	**186**	465	824	399	457
6	167	367	754	384	**150**	**112**	460	720	317	335
7	**378**	614	630	389	530	**104**	347	792	161	433
8	418	384	539	563	**275**	337	537	626	258	**208**
9	**57**	485	769	253	243	**32**	440	755	283	276
10	322	380	609	**302**	651	**127**	368	720	243	372
11	**219**	455	693	387	237	230	227	506	**206**	239
12	292	321	629	314	**224**	**172**	360	692	327	241
13	404	589	637	**319**	371	465	557	575	509	**424**
14	**167**	419	708	364	311	**46**	374	807	230	285
15	216	352	731	**177**	448	**73**	424	877	236	314
16	438	543	532	**349**	487	**180**	351	707	367	340
17	**133**	171	451	224	216	**46**	194	582	135	213
18	101	136	289	**98**	118	**18**	121	345	55	129

[a] II = information-integration model
[b] L = unidimensional model based on linewidth
[c] O = unidimensional model based on orientation
[d] C1 = first conjunctive model
[e] C2 = second conjunctive model

Smaller values indicate better model fit. The smallest values for each subject and task are highlighted.

Table B.1. *Experiment 2: BIC values (continued from previous page)*

	Third session (training)					*Fourth session (fMRI)*				
Subject	II[a]	L[b]	O[c]	C1[d]	C2[e]	II	L	O	C1	C2
1	**17**	401	910	253	350	**9**	435	820	231	330
2	**29**	353	841	253	313	**18**	414	823	245	503
3	**95**	374	680	239	294	**96**	381	809	252	307
4	319	342	435	**126**	483	334	243	496	**183**	485
5	**55**	425	689	239	362	**59**	413	751	315	272
6	**35**	405	761	279	610	**82**	388	771	281	286
7	**179**	455	752	322	345	**179**	401	794	276	404
8	**233**	384	504	261	336	**189**	500	616	220	350
9	**60**	334	771	232	305	**43**	357	872	236	316
10	**47**	402	834	277	305	**47**	344	758	266	259
11	**247**	275	592	270	356	**173**	289	653	284	300
12	**188**	435	725	376	273	**60**	438	866	276	199
13	**135**	403	814	299	283	**97**	500	830	247	580
14	**47**	405	767	217	670	**46**	379	860	301	252
15	**21**	405	851	206	335	**28**	406	859	299	292
16	**138**	343	656	224	344	**34**	392	770	250	244
17	**58**	251	693	201	170	**30**	397	714	309	328
18	**47**	177	413	66	171	**80**	115	298	96	138

[a] II = information-integration model
[b] L = unidimensional model based on linewidth
[c] O = unidimensional model based on orientation
[d] C1 = first conjunctive model
[e] C2 = second conjunctive model

Smaller values indicate better model fit. The smallest values for each subject and task are highlighted.

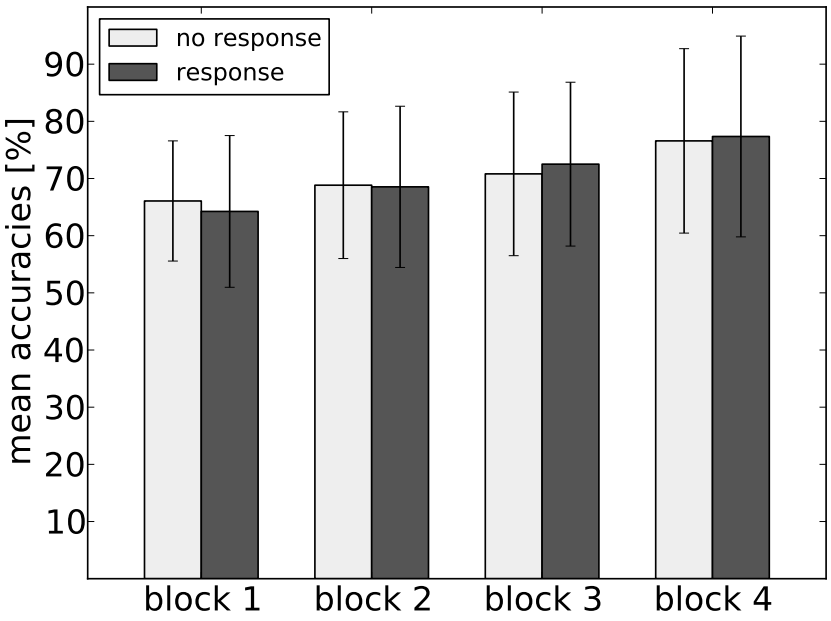

Figure B.1. *Experiment 2: Error rates in the pilot study.* Mean accuracies during the test trials of each block are depicted, error bars represent the standard deviation.

Dipl.-Psych. Reka Daniel-Weiner
Ostendstr. 21
70190 Stuttgart

Erklärung

hiermit erkläre ich, dass ich die von mir eingereichte Dissertation mit dem Thema

> *The influence of different forms of outcome information on the neural substrates of the acquisition and representation of categories*

selbständig verfasst, nicht schon als Dissertation verwendet habe und die benutzten Hilfsmittel und Quellen vollständig angegeben wurden.
Weiterhin erkläre ich, dass ich weder diese noch eine andere Arbeit zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.) an anderen Einrichtungen eingereicht habe.

Magdeburg, den 20. April 2012 Dipl.-Psych. Reka Daniel-Weiner