Cognitive control and binding in context-based decision-making : normal and dopamine deviant populations

Wouwe, N.C. van

Chapter 1

Introduction

Cognitive control

You are planning a dinner for close friends. Your intention is to compose something special so you search for a recipe and subsequently go to the grocery-shop to buy the ingredients. This is where things might go wrong: your memory is challenged to actively maintain the necessary ingredients while shopping and even if you make a shopping list, you might get distracted by products you are currently attracted to, like a chocolate cake, or the products you usually buy, like bread and fruit. In the current situation, the most optimal behavior to accomplish your goal (make a special dinner for friends) would require strong maintenance of relevant information (the recipe ingredients) without allowing any distraction from irrelevant incoming information (buy chocolate cake, fruit, bread) while shopping. Cognitive control refers to the process of adapting behavior to situational demands and overcome habitual or automatic behavior that is currently not optimal for performance. This may involve maintenance of task-relevant information, bias attention to task-relevant information, inhibit distracting information or update goals when necessary.

Modulation of control processes

Although healthy adults are generally well able to plan a dinner and buy the required ingredients, individual differences and challenging situational demands may limit the ability to successfully exert cognitive control. For example, individual differences in levels of the neurotransmitter dopamine (DA) affect control processes like maintenance

and updating of information and monitoring performance; these processes are impaired in DA deficient populations like patients with Parkinson's disease (PD), with attention hyperactivity disorder (ADHD) and schizophrenia (Braver, Barch, & Cohen, 1999; Frank, Santamaria, O'Reilly, & Willcutt, 2007; Moustafa, Cohen, Sherman, & Frank, 2008). Similarly, DA modulations by means of drug intake or rewarding feedback may alter control processes.

Moreover, drug studies (Beatty, 1995; Hyman & Nestler, 1993) have associated increased DA levels with an elevation of positive mood. Thus, it has been suggested that the impact of positive affect on cognition might be the result of a temporary increase of DA release in midbrain DA-generation centres, which is propagated to dopaminergic projection sites in other brain areas, such as the prefrontal cortex (Ashby, Isen, & Turken, 1999; Ashby, Valentin, & Turken, 2002). For example, several behavioural studies have indicated that transiently induced positive affect modulates control processes, generally leading to enhanced flexibility (Ashby et al., 1999; Dreisbach, 2006; Dreisbach & Goschke, 2004; Dreisbach et al., 2005; Kuhl & Kazén, 1999; Philips et al., 2002).

In addition to affective, or neurotransmitter, modulation of control, current task demands may affect control processes; for example the requirement to maintain task-relevant information during a temporally extended period or when strong habitual associations have to be overcome.

The following sections will address a number of influential theoretical frameworks that have been proposed to explain in what way these task demands may affect control processes in decision-making (section on context-based decision-making), the brain mechanisms involved in these control processes and current controversies between theoretical frameworks. The section on dopaminergic modulation of cognitive functions will concentrate on the theories that explain the role of DA in control processes. The studies that make up the current dissertation will be introduced related to these theoretical frameworks.

Development in models of control

Descriptive models of cognitive control first recognized the need for a goal-based control mechanism since habit- or routine-based behavior was insufficient to explain

adequate performance on novel tasks, or tasks that required flexible behavior (Baddeley, 1986; Norman & Shallice, 1986; Shiffrin & Schneider, 1977). In novel situations or situations where a habitual response pattern is inadequate, like driving to the office on a Saturday morning instead of going to the supermarket to buy groceries for a planned dinner, control has to be exerted to activate or inhibit task-relevant information. Baddeley described the capacities needed to exert this control; focus, divide, and switch attention. The control mechanism in the early descriptive models, which consisted of goal or task information, remained separate from storage and maintenance of stimulus information, whereas more recent models (Cohen, Dunbar, & McClelland, 1990; Kimberg & Farah, 1993) suggest that control and working memory (WM) functions are supported by the same mechanism. Indeed, in an updated version of Baddeley's classic model (2003) control and storage are both subserved by a central executive.

Descriptive models of WM and control have been formalized into computational models, allowing simulation studies to verify the models' assumptions and generate new hypotheses that can be tested empirically (e.g., Cohen, Braver, & Brown, 2002; Cohen, Dunbar, & McClelland, 1990; Kimberg & Farah, 1993). Kimberg and Farah simulated performance of frontal-lobe patients on a range of tasks: a Stroop task, Wisconsin Card Sorting Task, and a context-memory task. Although these tasks ask for different cognitive abilities, such as resolving response competition, decision-making based on feedback and maintenance of information in memory, impaired performance stemmed from damage in one single unit of the model; weakened associations between relevant stimulus and task information in WM. This model and subsequent computational models (Braver & Barch, 2002; Braver, Barch, & Cohen, 1999; Braver et al., 2001) depicted that control and WM function can be subserved by a single mechanism. Thus, in order to exert control, it is essential to represent, maintain, and update task-relevant information in WM that may bias attention during decision-making. If WM updating or maintenance is not intact, decision-making may be driven by automatically triggered stimulus-response (S-R) associations instead of by what is currently task-relevant.

But how does the system know when to exert control? Botvinick, Braver, Barch, Carter, and Cohen (2001) proposed a model that indicates the need for control by monitoring information processing for conflict. For instance, in a Stroop task a stimulus may activate two competing response tendencies, both word-color (red) and

word-name (blue). In this model, conflict is computed as the product of activation in competing response pathways and can be used to increase control on subsequent events. Evidence for a neural mechanism of control and models specifically relevant to the current work will be elaborated on in the section on context-based decision-making.

Context-based decision-making: stimulus- or goal-driven?

The models described in the previous section emphasize the importance of WM maintenance of rules, intentions, and other factors that make up the current task context when acting adaptively. Orthogonal to this goal-driven guidance, decisions may be biased also by stimulus-driven factors, such as the automatic reactivation of episodic associations that accompanied the current stimulus in a previous instance. This allows fast and automatic selection of a correct response in familiar, yet demanding environments. Both goal-driven and stimulus-driven biases may produce costs and benefits, depending for example on current task demands. Goal-driven and stimulus-driven factors in biased decision-making are typically investigated in separate studies, but may well apply simultaneously, sometimes converging on similar decisions but oftentimes yielding competing outcomes. Task information like goals and intentions as well as stimulus-based response or reward associations, are here referred to as the decision context.

Goal-driven guidance of behavior?

In the previously mentioned dinner example, the most optimal behavior to accomplish your goal (prepare a dinner for friends) would require strong maintenance of relevant information (the recipe ingredients). However, if you receive new information (e.g., you learn that one of your friends is a vegetarian), you have to adapt your plans, change the recipe and buy other ingredients. Updating your current intentions is crucial in this case, whereas maintenance of the previous goal would be erroneous.

Behavioral costs of deficits in proactive preparation, like deficits in WM maintenance and updating, have been revealed by classic paradigms like Stroop, Eriksen

Flanker task (Eriksen & Eriksen 1974; Stroop, 1935), set shifting (Jersild, 1927), Wisconsin Card Sorting test (Grant & Berg, 1948; Milner, 1963) and the *AX* version of the Continuous Performance Task (*AX*-CPT; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956). The *AX*-CPT will be focus of the studies in chapter 2 to 4, details regarding this paradigm can be found in these chapters.

The inability to maintain task-relevant information or the inability to update information leads to distractibility or perseveration costs. For example, in a Stroop task participants have to maintain task goal information 'name the color of the word', and refrain from a more habitual response tendency to read the word, in order to perform the task correctly. Switch costs in task switching paradigms can be explained similarly. During task switching subjects constantly have to update relevant task information (indicated by the task cue), which takes effort and time. These switch costs are reduced when subjects are given time to prepare for the upcoming task (Meiran, 1996; Meiran, Chorev, Sapir, 2000; Rogers & Monsell, 1995).

Braver and colleagues (Braver & Barch, 2002; Braver, Satpute, Rush, Racine, & Barch, 2005) described goal-driven cognitive control in the *AX*-CPT in terms of a Context Processing Model. They used the term context representation to refer to the goal or stimulus representations that influence planning, behavior and attentional processes. Top-down control can be exerted because the context information biases or primes the activation of a response or goal, as previously connected with that particular context information, in subsequent trials. On each *AX*-CPT trial participants are presented with a cue stimulus and a subsequent probe stimulus, and are instructed to respond if a target probe (*X*) is immediately preceded by a specific cue (*A*) but to refrain from responding in all other sequences (*AY, BX* or *BY*). Target trials (*AX*) typically occur on the majority of trials in the *AX*-CPT task; this frequency induces a strong bias to issue a target response, even when either the cue (*BX*) or the probe (*AY*) designates that a response has to be withheld. Successful performance on *BX* trials is often interpreted as a result of top-down control. According to the context processing model, performance costs occur if a context primes an incorrect response, for example in *AY* trials, or if subjects fail to maintain the relevant context information (Braver et al., 2001).

Neuroimaging studies (Sakai & Passingham, 2003), animal research (Wallis, Anderson, & Miller, 2001), and neural-network simulations (O'Reilly, Noelle, Braver, & Cohen, 2002) suggest that the representation, maintenance, and updating of rule and

task information is regulated by the Prefrontal Cortex (PFC). The PFC is part of a densely connected circuit; PFC receives input from sensory cortices, projects to motor cortices and is connected with subcortical regions like the basal ganglia (BG), amygdala and hippocampus. In addition to that, most regions within PFC are highly interconnected (Miller & Cohen, 2001). Together, these structures make up the necessary infrastructure to integrate different stimulus features, to activate rules and to plan for future action (Duncan, 2001; Miller & Cohen, 2001). More specifically, imaging studies have shown activation in the dorsolateral part of the PFC (DLPFC) with increase in WM demand (Cohen et al., 1997), selection between competing responses (Bunge, Hazeltine, Scanlone, Rosen, & Gabrieli, 2002; Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000) and allocation of attention (MacDonald, Cohen, Stenger, & Carter, 2000).

In sum, DLPFC is thought to apply control by maintenance and updating of task-relevant information against interfering information from the environment and by contributing more attention to task-relevant information. Signals from DLPFC to visual or associative brain areas may increase activation for task-relevant stimulus properties (Botvinick et al., 2001; Desimone & Duncan, 1995; Dove, Pollman, Schubert, Wiggins, & von Cramon, 2000; Egner & Hirsch, 2005; Gruber, Karch, Schlueter, Falkai, & Goschke, 2006; Kerns et al., 2004; MacDonald, Cohen, Stenger, & Carter, 2000).

Although task performance significantly benefits from proactive task preparation, decision-making also requires a more ad-hoc control mechanism which is responsive to new incoming stimulus information or indicates the need for control on future trials.

Whereas proactive control is generally attributed to the PFC (Boettiger & D'Esposito, 2005), the anterior cingulate cortex (ACC) is thought to be engaged in the reactive control process; i.e. detecting possible interference between co-activated responses or erroneous response tendencies and signaling the PFC for need for control on subsequent trials. Botvinick et al. (Botvinick, Cohen & Carter, 2004; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Botvinick et al., 2001) proposed the conflict-monitoring theory of the ACC which is corroborated by a large body of imaging studies. Activity in the ACC is associated with conflict between competing stimulus or response tendencies in several paradigms like the Stroop task (Barch et al., 2001), the Eriksen flanker task (Botvinick et al., 1999; Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002; Casey et al., 2000; Durston et al., 2003) and the Simon task (Kerns, 2006; Peterson et

al., 2002). An event-related potential (ERP) component associated with conflict, the N2 (a negative deflection that emerges around 150-300 ms after stimulus presentation), also increases with high-conflict trials and has been localized in the ACC (Nieuwenhuis, Yeung, Ridderinkhof, & van den Wildenberg, 2003; van Veen & Carter, 2002a; 2002b; Yeung & Cohen, 2006). Activity in ACC subsequently triggers the DLPFC to resolve the conflict (Durston et al., 2003; Kerns et al., 2004; MacDonald et al., 2000) and enhance control on subsequent trials or the current trial (DePisapia & Braver, 2004).

Other explanations with respect to role of ACC in performance monitoring incorporate the DA system (error detection and reinforcement learning) and are introduced in the section on dopaminergic modulation of cognitive functions. In chapter 3 and chapter 4 monitoring-related brain activation in the ACC is investigated by means of an ERP and functional magnetic resonance imaging (fMRI) study.

Association-driven guidance of behavior?

In daily life routine activities can be performed quickly and without much effort, which is often beneficial. For example, while driving home you can solve a theoretical problem or plan a dinner meal. However, automatic behavior may also go along with costs in situations when incorrect behavior is automatically triggered by the environment, for example when entering the car this may automatically trigger driving home whereas the current intention was to go to a restaurant.

Different theoretical accounts have addressed the question how associations between stimuli and actions may be established and when these associations bias performance. To begin with, associations between stimulus and response features may stem from attention or contiguity: attending to an event creates an episodic memory trace in which stimulus and response features of an event are bound together (Hommel, 1998; Hommel, Proctor, & Vu, 2004). On future occasions re-occurrence of the same stimulus may activate the previously associated motor plan (Hommel, 2004; Logan, 1988). The feature integration account was originally motivated by the binding problem; when confronted with multiple objects, how does the brain 'know' which perceptual features belong together in one object (Treisman & Gelade, 1980). Kahneman, Treisman and Gibbs (1992) proposed that representations of different

perceptual features that are distributed over the brain but belong to the same object, can become integrated in an 'object file'. This object file is an episodic trace containing information about the relationship between object-features. Hommel (1998, 2004) extended this object-file account of visual feature integration to the integration of perception and action features of an event.

Moreover, S-R associations may also be established by reinforcement learning and strengthened with extensive practice (cf. Daw, Niv, & Dayan, 2005; Balleine & Dickinson, 1998). For example, if an action in response to a certain event is positively rewarded, the association between this action and the event is more likely to be strengthened because it enables maximizing reward in the future. Similarly, when an action has habitually been performed with a specific stimulus, their association strengthens. The aim of the present studies (i.e. chapter 2 and 3) was not to determine the conditions of when or how these bindings can be established, but to investigate whether an association-driven bias, in addition to a goal-driven bias, can explain performance costs in decision-making and to determine the neural correlates that explain these costs. Chapter 5 and 6, on the other hand, did take a closer look at the neural requisites that facilitate learning stimulus-action associations by means of reward.

Biases derived from episodic retrieval can lead to performance costs, for example when features of the current event partially correspond with features of a recent episodic memory trace, such as when a stimulus was associated with one action in a previous instance but currently designates a different action. According to binding accounts, performance benefits in situations where stimulus and action features are all repeated or all changed (Hommel et al., 2004; Mayr, Awh, & Laurey, 2003; Pashler & Baylis, 1991), whereas performance is hampered on partial-repetition trials (Hommel et al., 2004; Notebaert, Soetens, & Melis, 2001) either because previous associations need to be "undone" or because the conflict they induce needs to be overcome. Moreover, binding accounts might explain, at least in part, performance costs previously attributed exclusively to top-down control operations. Studies on Flanker interference (Mayr et al., 2003), negative priming (Huang, Holcombe, & Pashler, 2004; Tipper, 2001), inhibition (Verbruggen, Logan, Liefooghe, & Vandierendonck, 2008), task switching (Waszak, Hommel, & Allport, 2004), and spatial incompatibility (Hommel et al., 2004) indicated that the reactivation of competing information as a result of a

retrieved episodic binding created on a previous trial bias performance on the current trial.

Animal associative-learning studies and computational models based on animal work also pointed out the behavioral costs of S-R associations (Daw et al., 2005). Although the associations built up in these experiments are in many aspects different from the binding effects studied in humans (i.e. they are gradually learnt by extensive practice and oftentimes established by rewarding feedback), they point to similar performance costs with decision-making; it takes time to undo or adapt previously learnt associations. For example, if an outcome to a certain stimulus or action changes, this cannot be immediately re-evaluated and adapted, but it takes several trials to build up a new association. Likewise, it takes more time on the current trial to undo a previous association or rebuild a new association (as found in human behavioral studies of binding).

The neural representation of episodic bindings and the neural correlates of associated performance costs or benefits recently received some attention. A neuroimaging study by O'Craven, Downing, and Kanwisher (1999) showed that attending to an event results in the cross-referencing and co-activation of the relevant and irrelevant features of this event, as reflected by brain activation in the corresponding brain areas. Follow-up work of Keizer et al. (2008) showed that this co-activation of object features indeed creates bindings between them; repeating one feature increases the activation of brain areas coding for the previously bound associated features. It remains unclear however, whether traditional 'top-down control' areas, like DLPFC and ACC, will also be recruited with the reactivation of competing information as a result of a retrieved episodic binding. The fMRI study in chapter 3 will provide insight into the way binding-related performance costs or benefits may be represented in the brain. However, these results do not inform us about the neural mechanism that established these bindings.

Currently, there are several neurobiologically plausible mechanisms that can account for binding. First, evidence from single-cell recordings in rats (Engel & Singer, 2001) and magnetoencephalography (MEG) and electroencephalography (EEG) studies with humans (Tallon-Baudry & Bertrand, 1999) indicated that temporal synchronization of gamma oscillations binds properties of an object with several sensory modalities into an integrated object representation. Second, imaging studies have shown that especially

the Premotor Cortex, including the Supplementary Motor Area (SMA), could serve as a facilitator for the integration of action planning and sensorimotor data (cf. Grezes & Decety, 2001; Jackson & Decety, 2004; Schubotz & Von Cramon, 2003). Third, several investigations suggest that visuomotor integration is driven by the dopaminergic system. Colzato, van Wouwe, & Hommel (2007a) recently showed that binding of visual and action features is modulated by the presentation of affect-inducing pictures, which can be assumed to stimulate the dopaminergic system. Additionally, (Colzato, et al., 2007a; Colzato, van Wouwe, & Hommel, 2007b) found evidence that the spontaneous eye-blink rate (a functional marker of central dopaminergic function) reliably predicts the strength of the binding between task-relevant stimulus features and the response.

The section on dopaminergic control of cognitive functions will provide a more detailed account of the brain areas involved in learning reward-based S-R associations.

Stimulus- and goal-driven biases in decision-making?

The previous sections illustrated the importance of either stimulus- or goal-driven biases of behavior. Each account has its own advantages and disadvantages in determining action control. Habitual behavior is not as flexible as proactive behavior, but allows fast and automatic selection of a correct response in familiar, yet demanding environments, like changing gears in busy traffic. Proactive goal-directed behavior allows on the fly behavioral adaptations when the circumstances change, but is also costly because memory is highly taxed; i.e. goals and predictions about immediate consequences of an action have to be kept in mind.

The contribution of both perspectives can be pointed out by how they account for the Gratton effect (Gratton, Coles, & Donchin, 1992). The Gratton effect consists of a reduced congruency effect (i.e. reduced performance costs on incongruent relative to congruent trials) after incongruent compared to congruent trials in an Eriksen Flanker task. This effect has also been replicated in Simon (Stűrmer, Leuthold, Soetens, Schroter, & Sommer, 2002) and Stroop tasks (Kerns et al., 2004). More specifically, in the Flanker task reaction times on incongruent trials subsequent to incongruent trials are faster than incongruent trials succeeding congruent trials. Likewise, performance is

faster on congruent trials preceded by a congruent trial than on congruent trials preceded by an incongruent trial. A typical goal-driven account like the conflict-monitoring theory (Botvinick et al., 2001) suggests that conflict is detected between competing response tendencies in incongruent trials, which increases control and improves performance on the next trial.

Others, however, have reported that the conflict adaptation effect in a Flanker task is due to the repetition of specific stimulus episodes and does not necessarily involve control such as monitoring response conflict (Mayr et al., 2003; Nieuwenhuis et al., 2006). In a Flanker task, stimulus-specific repetitions are especially present in congruent followed by congruent trials or incongruent followed by incongruent trials, thus performance on these trials may be fast and provide an alternative explanation for the conflict adaptation effect.

Alternatively, several studies point towards an integration between stimulus-driven and goal-driven accounts because they have been shown to concurrently contribute to performance in a number of behavioral paradigms (for a review see Egner, 2007; Notebaert, Gevers, Verbruggen, & Liefooghe, 2006; Ullsperger, Bylsma, & Botnivick, 2005; Verguts & Notebaert, 2008). Verguts and Notebaert (2008) proposed a model that integrates conflict-based adjustments in control with association-based effects. Associations between stimuli and responses might be used by a conflict-monitoring system to indicate the need for control on future trials. That is, if both stimulus and action features remain equal between the current and the previous event there is no conflict signal and thus no need for increased control. Other studies also indicate that conflict-driven control may be applied at different levels, from global task level to item-specific control (Blais, Robidoux, Risko, & Besner, 2007; Egner, 2008).

Chapter 2 and 3 investigated the role of goal-driven versus stimulus-driven processes in context decision-making using an *AX* version of the continuous performance task (*AX*-CPT). Target trials (*AX*) typically occur on the majority of trials in the *AX*-CPT task; this frequency induces a strong bias to issue a target response, even when either the cue (*BX*) or the probe (*AY*) designates that a response has to be withheld. Goal-driven preparation in this task may thus lead to performance costs on *AY* trials (cue indicates incorrect target response) and benefits on *EX* trials (cue correctly indicates nontarget response). Episodic bindings in the *AX*-CPT may arise

between specific cue and probe stimuli and reactivation of competing information as a result of these episodic bindings may bias performance on subsequent trials.

Two behavioral experiments tested the contribution of S-R repetitions vis-à-vis top-down control (chapter 2) and an fMRI study sought to test the relative importance of stimulus-stimulus (S-S) repetitions vis-à-vis top-down control (chapter 3) in context-based decision-making. These studies were inspired by human behavioral studies on stimulus-driven versus goal-driven accounts. Note that the effect of episodic bindings on a global task level was examined instead of trial-by-trial effects as in some other studies (cf. Hommel et al., 2004).

Dopaminergic modulation of cognitive functions

DA projections in the brain

DA neurons project to a wide range of cortical areas and thus DA seems a good candidate to modulate several aspects of decision-making. Dopaminergic projections in the brain can be broadly divided into the mesocorticolimbic and nigrostriatal system (Ashby et al., 1999). The first consists of DA neurons projecting from the ventral tegmental area to brain areas such as the ACC, the DLPFC, hippocampus and amygdala and the latter consists of DA projections from the substantia nigra pars compacta to the striatum.

The role of DA in cognitive control has been investigated by a broad range of studies, such as drug studies in healthy subjects (Barch, 2004; Frank & O'Reilly, 2006; Kimberg & D'Esposito, 2003), DA-deficient patient populations such as PD, ADHD and schizophrenia (Braver et al., 1999; Frank et al., 2007; Moustafa et al., 2008), genetic variation in DA activity (Mattay et al., 2003; Nolan, Bilder, Lachman, & Volavka, 2004) or indirect DA manipulations by means of reward or positive affect induction (Ashby & Turken, 1999). Chapter 4 addressed the modulating influence of positive affect on control processes in context-based decision-making.

Whereas chapter 2-4 sought to investigate the contribution of goal-driven and stimulus-driven biases in context-based decision-making (performance costs that result

from different types of context), chapter 5 and 6 studied context formation, that is, the formation of stimulus-action-reward associations and the effect of BG modulations on this learning process. More specifically, chapter 5 examined the effect of DA modulations in context formation related to substructures in the striatum. Therefore, theoretical accounts with respect to the role of DA in control processes involved in context-based decision-making will be discussed, for instance, the role of DA in preparatory proactive control processes and in subsequent evaluative processes. Chapter 4 will elaborate on evaluative control and the DA theory of positive affect and chapter 5 will discuss the role of DA in the context formation, i.e. stimulus-action-reward associations. Chapter 6 studied the effect of a more specific BG modulation (deep-brain stimulation of the subthalamic nucleus) in context formation, therefore some additional background will be provided on BG structure and function.

Dopaminergic modulation of cognitive control processes

DA and proactive preparation

Computational modeling work of Braver and Cohen (2000) and Cohen, Braver, and Brown (2002) postulated that tonic DA in the PFC regulates the active maintenance of task-relevant information. Phasic increases of DA in prefrontal cortex, elicited by reward-predicting stimuli, serve as a gating signal to update WM, replacing the current representation with one that will guide behavior to reward (Braver & Cohen, 2000; Durstewitz, Seamans & Sejnowksi, 2000; Holroyd & Coles, 2002). Several lines of evidence support the link between proactive preparation and DA function. Deficient memory maintenance or updating, due to a DA depletion, impacts performance on tasks that require adaptive behavior. Braver demonstrated that populations with a decline in DA functioning in PFC, like older adults, are impaired in proactive control of behavior on the *AX*-CPT task (Braver et al., 2001). Systematic manipulation of DA in humans by using DA agonists (for an overview see Barch, 2004) pointed to enhancing effects of DA on WM tasks such as the spatial delayed match to sample test and the N-Back task, although questions remain about the role of dosage, individual differences (i.e. in WM capacity and baseline DA) and task characteristics (Kimberg & D'Esposito, 2003). Generally, the effect of DA on WM is suggested to follow a U-shaped curve. This

entails that there is an optimal level of dopaminergic activity for intact WM performance and that hypo or hyper-dopaminergic states may lead to impairments (Arnsten & Goldman-Rakic, 1998).

Recently, it has been posed that whereas DA in the PFC enables stable maintenance of information, DA in the BG modulates updating WM representations (Frank & O'Reilly, 2006; Cools, Sheridan, Jacobs, & D'Esposito, 2007; McNab & Klingberg 2008; Moustafa et al., 2008). Computational models of basal ganglia-frontal cortical systems indicate that a mechanism similar to reinforcement learning may subserve WM-updating functions (Frank & O'Reilly, 2006; Moustafa et al., 2008; O'Reilly & Frank, 2006). Feedback-based phasic DA changes in the BG facilitate learning which information is relevant and has to be updated into prefrontal cortex on future events and which information has to be suppressed in the future. The modulatory effect of DA on projections from the BG to the PFC may enable rapid and selective updating compared to the slower and more diffuse effects of direct projections of DA to the PFC (O'Reilly, 2006; Seamans & Yang, 2004). A more detailed description of the learning mechanism will be provided in the section about DA in reinforcement learning.

DA and evaluative control: performance monitoring

Whereas some argue that the ACC is associated with conflict-detection between competing stimulus or response tendencies (Botvinick et al., 2001), others suggest it is involved in error-detection. According to Holroyd and Coles (2002), the evaluative function of the ACC is affected by DA fluctuations in the nigrostriatal DA system (i.e. the striatum is connected with the ACC in a basal-ganglia-thalamo-cortical circuit, (Alexander, DeLong, & Strick, 1986; Alexander, Crutcher, & DeLong, 1990). The ACC is thought to detect discrepancies between expected and actual outcomes of behavior and thus monitors performance for unfavorable outcomes or errors (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). When outcomes are worse than expected, this coincides with a phasic dip of DA in the mesencephalic DA system which disinhibits the ACC (Holroyd & Coles, 2002, but see also Botvinick, 2007; Frank, D'Lauro, & Curran, 2007; Jocham & Ullsperger, 2009, for an opposing temporal account on the ACC and DA signals). These dips in DA may thereby drive NoGo learning to avoid selecting the same response (in reaction to a stimulus) in the future while DA bursts support Go learning to maximize reward (Frank, Woroch & Curran, 2005).

The findings on error-induced activity in the ACC are supported by evidence from ERP studies (Gehring, Goss, Coles, Meyer, & Donchin, 1993; Falkenstein et al. 1991), especially on the error-related negativity (ERN), thought to be generated by the ACC (Dehaene, Posner, & Tucker, 1994; van Veen & Carter, 2002a). fMRI studies also found increased ACC activity with errors (see van Veen, Holroyd, Cohen, Stenger, & Carter, 2004) or negative feedback (Holroyd et al., 2004; Knutson et al., 2000; Ullsperger & von Cramon, 2003).

Experimental manipulations of DA on performance monitoring are mainly conducted by studying the ERN in DA-deficient patients, in genetic variations of DA receptors or after pharmacological manipulations (for a review, see Jocham & Ullsperger, 2009). The ERN is a medial-frontal deflection in the ERP, peaking around 100 ms after an erroneous response (Falkenstein, Hohnsbein, Hoorman, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). A sizeable ERN is thought to reflect a phasic dip in DA when outcomes are worse than expected (but see also Berridge, 2007; Redgrave & Gurney, 2006).

The ERN is relatively spared in patients with PD off medication, who have low striatal DA levels (Holroyd et al., 2002), but diminished in populations with high levels of DA, like PD patients on medication (L-Dopa) and patients with schizophrenia (Bates, Liddle, Kiehl, & Ngan, 2004; Falkenstein et al., 2001; Ito & Kitagawa, 2006; Stemmer, Segalowitz, Dywan, Panisset, & Melmed, 2007). Simulations of a variety of DA-related effects by Frank et al. (2004; 2005) have indicated that a smaller ERN is elicited in the same conditions where learning from negative feedback is diminished. Surprisingly, administration of a D2-antagonist haloperidol (thought to reduce DA) also reduced the ERN (de Bruijn, 2006; Zirnheld, 2004) and improved learning from positive but not from negative feedback (Frank & O'Reilly, 2006). Jocham and Ullsperger (2009) suggest that inhibiting D2 receptors actually increases DA concentrations (by blocking autoreceptors). The reduction of the ERN with L-DOPA and pharmacological D2 receptor blockade might thus also be due to enhanced activity at D1 receptors.

Both the error-based (Holroyd & Coles, 2002) and conflict-based theory (Botvinick et al., 2001) of performance monitoring explain ACC activity in terms of competing processes. However, the conflict theory of performance monitoring accounts for conflict in error and correct trials, but it does not explain feedback processing. On the other hand, the error-monitoring theory accounts for error processing without conflict,

but not for conflict in correct trials. Moreover, error-monitoring theory is based on reinforcement-learning theories (Schultz, 2002) which assume that learning and feedback processing is modulated by phasic changes in DA.

Yeung, Botvinick, and Cohen (2004) suggested, based on a series of simulations that the conflict- and error-related ERP both reflect monitoring response conflict. The ERN reflects post-response conflict that develops following errors because continued stimulus processing leads to post-error activation of the correct response which conflicts with the produced incorrect response.

Verguts and Notebaert (2008) proposed a Hebbian learning model that integrates the evaluative acccunt of the ACC (i.e. error-based changes in behavior) with conflict monitoring (conflict-based changes in behavior). They pointed out that like errors, conflict signals from ACC to substantia nigra can be used as learning signal to adapt behavior on future events.

Basal Ganglia modulations and reward-based learning

BG structure and function: striatum and subthalamic nucleus

The BG consist of the striatum, the globus pallidus, the substantia nigra and the subthalamic nucleus (STN). The striatum is the main input structure and receives projections from the cortex and subcortial areas, i.e. dopaminergic projections from the substantia nigra and the ventral tegmental area. The striatum is connected with cortical brain areas in multiple cortico-striatal loops (Alexander et al., 1986; Alexander et al., 1990);) the motor, oculomotor, orbitofrontal, dorsolateral and anterior cingulate loop, subserving motor control, cognitive and motivational behavior. Because each of these loops serves a different function, the striatum has been posed to integrate motivational, cognitive and motor control information. More specifically, the dorsal part of the striatum can be divided into the caudate nucleus and the putamen. The putamen and caudate nucleus are located in different corticostriatal loops; the putamen is embedded in the motor loop while the caudate nucleus is connected to the lateral orbitofrontal cortex and the dorsolateral prefrontal cortex (Alexander et al., 1986; Alexander et al., 1990). The ventral part of the striatum primarily consists of the nucleus accumbens,

although ventral putamen and caudate are also considered being part of the ventral striatum (Berns et al., 2001; Seger & Cincotta, 2005).

To independently facilitate execution and inhibition of motor commands in the cortex, the BG consist of two pathways projecting from the striatum through different nuclei to the thalamus and up to the cortex (Alexander, Crutcher & DeLong, 1990; Mink, 1996; but see also Kawaguchi, Wilson & Emson, 1990; Levesque & Parent 2005, questioning the simplicity of this model); the direct Go pathway and the indirect NoGo pathway. Striatal Go cells disinhibit the thalamus via inhibition of the internal segment of the Globus Pallidus (GPi). The thalamus in its turn facilitates the execution of an action that is currently represented in the cortex. Striatal NoGo cells inhibit the thalamus indirectly via inhibitory projections to the external segment of the Globus Pallidus (GPe) and the GPi, which suppresses action execution. DA differentially affects these pathways (Aubert, Ghorayeb, Normand, & Bloch, 2000; Gerfen, 1992; Hernandez-Lopez, Bargas, Surmeier, Reyes, & Galarraga, 1997; Hernandez-Lopez et al., 2000); it activates D1 receptor cells (mostly in the Go pathway), whereas it inhibits D2 receptor cells (mostly in the NoGo pathway).

Similar to activation and inhibition of action plans, parallel BG circuits may be engaged in modulation of other cortical areas, such as updating WM representations in PFC (Beiser & Houk, 1998; Frank, Loughry, & O'Reilly, 2001; O'Reilly & Frank, 2006).

The STN is embedded in the indirect NoGo pathway as well as in a hyperdirect pathway (Orieux, Francois, Feger, & Hirsch, 2002). The STN receives input from GPe and targets the GPi with excitatory projections and inhibits the thalamus and thereby inhibits action execution. In the hyperdirect pathway projections from the cortex directly target the STN which subsequently activates the GPi and thereby bypasses the striatum. Thus, if multiple competing responses are activated in the motor cortex, the STN gets increasingly excited, thus preventing a premature response (stronger NoGo signal).

Additionally, STN receives sensorimotor, cognitive and limbic input from the GPe that are functionally separate. However, the boundaries between these territories within the STN are not sharply defined (Karachi et al., 2005). Nor is there a clear segregation between modalities in the output of the STN (Sato, Parent, Levesque, & Parent, 2000). Mallet et al. (2007) recently proposed that the STN not only regulates

input from different modalities, but also integrates sensorimotor, emotional and cognitive aspects of behavior.

DA and reward-based decision-learning

Although the BG are traditionally known to contribute to motor function, (Alexander et al., 1986; Alexander et al., 1990), the contemporary view suggests involvement of the BG in several types of learning, including habit formation, procedural skill learning, and reward-based decision-learning (Brown & Marsden, 1998; Kimura, 1995; Knowlton, Mangels, & Squire, 1996; Packard & Knowlton, 2002; Schultz, Tremblay, & Hollerman, 2003).

Several lines of research (with DA-deficient populations, human drug studies, animal studies and computational modeling) have indicated that DA, via projections from the substantia nigra and ventral tegmental area innervating the striatum, plays a modulatory role in learning (Arnsten, 1998; Cools, 2006; Daw et al., 2005; Eyny & Horvitz, 2003; Frank, 2005; O'Reilly & Frank, 2006; Schultz, 2002). When nonhuman primates perform a stimulus-reward task, midbrain DA neurons initially fire with reward delivery (Schultz, 2002). Positron emission tomography (PET) studies in humans also revealed that DA in the striatum increased when receiving unexpected reward (Koepp et al., 1998; Pappata et al., 2002). During the learning process, this activity shifts in time from the actual reward to the presentation of a stimulus that predicts the reward (Schultz, 2002) and links between stimuli and responses are adjusted to minimize error between predicted and actual outcomes (the reward-prediction error). These prediction errors are coded by changes in firing rate of the dopaminergic neurons.

A useful explanation of how DA might modulate the BG and reward-based decision-learning was provided by Frank (2005; Frank & O'Reilly, 2006; O'Reilly & Frank, 2006) who proposed a mechanistic account to explain the modulatory role of phasic changes in DA with reward-based decision-learning. When outcomes are worse than expected, this coincides with a phasic dip of DA in the striatum, which removes the inhibitory effect of DA on a specific type of DA receptors, the D2 receptor cells (Frank, 2005). Activity in D2 receptor cells, found most prominently in the NoGo pathway, has an inhibitory effect on the thalamus and thereby suppresses actions from being executed. Dips in DA may thus drive NoGo learning to avoid selecting the same response (with a

specific stimulus) in the future. Phasic DA bursts on the other hand, induced by outcomes that are better than expected, activate another type of DA receptors, D1 receptor cells (found primarily in the Go pathway), which disinhibit the thalamus and thereby facilitate the execution of an action. This supports Go learning to maximize reward (Frank, 2005).

This model is corroborated by findings from DA-deficient populations such as PD and ADHD (Cools et al., 2001; Frank, Santamaria, O'Reilly, & Willcutt, 2007; Shohamy, Myers, Geghman, Sage, & Gluck, 2005; Swainson et al., 2000) and drug studies in humans and animals (Frank & O'Reilly, 2006; Goto & Grace, 2005); DA medication or drugs that continuously activate D2 receptors block the effects of DA dips which impair No-Go learning, whereas Go-learning may remains intact. A lack of DA, for example in PD patients off medication, reverses this bias. Although this model points out the way in which DA affects reward-based decision-learning, it does not specify whether DA might differentially modulate substructures within the striatum, such as the caudate and putamen, and the functions related to these structures.

The caudate and putamen have recently been functionally dissociated in an fMRI study by Haruno and Kawato (2006a; 2006b). They studied the activity of putamen and caudate nucleus during the performance of a probabilistic-learning task. This task measured the evolvement of expectancies about the outcomes of a decision as subjects attempted to learn novel stimulus-action-reward associations. Two aspects of learning can be computed from this task by means of the Q-learning algorithm (Sutton & Barto, 1998), which is an implementation of a temporal difference model. First, a reward prediction error (RPE) value, which was used to infer how proficient subjects were at using errors between anticipated rewards and actual rewards as a basis for adjusting decision-making on future trials. Haruno and Kawato reasoned that this value is particularly meaningful in the early stages of learning when subjects have little experience with which actions maximize rewards and consequently make a higher number of unrewarded decisions. Second, a stimulus-action-dependent reward prediction (SADRP) value was computed to capture the proficiency of learning which response maximized reward for each stimulus. Higher SADRP values reflect more effective learning of stimulus-action-reward associations, and hence, are maximal at the later stages of the task.

The results showed that caudate nucleus activity corresponded closely to the processing of reward, or more particularly the RPE value, during the early stages of learning. In contrast, putamen activity corresponded most closely to the successful formation of specific SADRP associations that became evident at the later stages of learning. These functional patterns fit well with the proposal that the putamen is embedded in corticostriatal motor loops that are responsible for mapping stimuli to responses and the proposal that the caudate nucleus forms functional loops with lateral orbitofrontal and dorsolateral prefrontal cortices, areas that are tied to complex cognitive and affective aspects of processing.

Chapter 5 used a similar probabilistic learning task to test the effect of a dopaminergic modulation on reward-based decision-learning processes linked to caudate and putamen, by means of a PD ON-OFF medication study.

Relevant to the investigation of chapter 5 are findings that the dopaminergic projections to regions of the striatum are affected differentially by the progression of the disease. Early in the disease, DA is more severely depleted in the motor (including putamen and supplementary motor areas) and dorsolateral loops (including the DLPFC and the dorsolateral head of the caudate) compared to the orbitofrontal (lateral OFC, ventromedial head of caudate) and anterior cingulate (anterior cingulate, ventral striatum) loops (Kaasinen & Rinne, 2002). The primary treatment of PD aims to increase DA availability and activity, including medications functioning as a DA precursor or as a DA agonist (Hornykiewicz, 1974). DA medication clearly enhances cognitive and motor functions supported by the severely depleted dorsal striatum (dorsal caudate & putamen), whereas studies that investigated DA-induced improvement of functions related to the ventral striatum, such as reward processing, show more ambiguous results (Cools, Barker, Sahakian, & Robbins, 2001; Czernecki et al., 2002; Frank, 2005; Gotham, 1988; Shohamy et al., 2005; Swainson et al., 2000). A more detailed background on PD, DA and reward-based decision-learning is provided in chapter 5.

We predicted that DA medication would have the strongest beneficial effect on the formation of stimulus-action-reward associations (putamen). We were less confident in our predictions about RPE values (associated with caudate and ventral striatum), since the caudate is thought to be less affected by PD than the putamen.

STN and reward-based decision-learning

Recently, modulations of the STN by deep brain stimulation (DBS), an alternative treatment used for PD, have also contributed to our understanding of the role of the BG in action control. In a DA-depleted brain, such as in PD, D1 receptors in the direct and D2 receptors in the indirect pathways are not effectively modulated by DA. Low levels of DA result in excessive activity in D2 NoGo striatal neurons. Additionally, these low levels of DA indirectly (via the GPe) remove the inhibition from the STN, resulting in an overactive STN and thereby an even stronger NoGo signal. Although there is an ongoing debate about the specific mechanisms underlying the therapeutic effect of STN stimulation (Benazzouz & Hallett, 2000; Bergman, Wichmann, & DeLong, 1990; Liu, Postupna, Falkenberg, & Anderson, 2006; Meissner et al., 2005), DBS is currently thought to inactivate the STN. This removes excessive activation of the GPi and thus disinhibits the thalamus, thereby facilitating thalamic excitation of the cortex.

Substantial evidence from animal studies (Baunez et al., 2001; Bergman, Wichmann, Karmon,& DeLong, 1994; Karachi et al., 2005) and PD patient studies (Jahanshahi al., 2000; Schroeder et al., 2002; Witt et al., 2004; van den Wildenberg et al., 2006) showed that the STN is critically involved in both motor control and action selection (Boraud, Bezard, Bioulac, & Gross, 2002).

Similar to the input from premotor cortex, the STN may receive reward-related information (i.e. expected magnitude of reward) from medial OFC projections to STN, as has been shown in rats (Maurice et al., 1998), and holding the response output system (GPi; thalamus) in check until the expected reward options for a certain response are evaluated. Stimulating the STN may disinhibit the limbic circuits analogous to the disinhibition of motor circuits. The electrodes used to stimulate the STN are usually implemented in the motor areas of STN (part of the corticostriatal motor loop) to treat PD related motor impairments. However, according to Mallet et al. (2007), stimulating the motor area of the STN may affect cognitive and limbic loops as well, because STN output is not sharply segregated.

However, the role of the STN and the effects of DBS of STN on reward-based decision-learning processes have been studied in a limited number of studies and the effects of DBS of STN are variable; both impairments and improvements in reward based learning are found (El Massoui et al., 2007; Frank et al., 2007; Funkewiez et al., 2006).

Chapter 6 tested the effect of a STN modulation of reward-based decision-learning processes linked to the caudate and putamen, by means of a PD ON-OFF STN stimulation study. Since STN stimulation clearly enhances motor functions (Kleiner-Fisman et al., 2003), which are primarily supported by regions of the dorsal putamen, we predicted that STN stimulation would have a beneficial effect on the formation of stimulus-action-reward associations (associated with the putamen). Given limited and contrasting findings regarding the role of STN and the effects of STN DBS on reward processing, the effects of STN DBS on RPE processing (associated with the caudate) are more explorative.

Outline of the dissertation

The dissertation focused on two general questions: Do goal-driven processes as well as stimulus-driven processes contribute to context-based decision-making? In what way do modulations of brain areas innervated by DA, i.e. striatum or prefrontal cortex, affect control processes in context-based decision-making and affect the acquisition of S-R associations? Both questions can be divided into several subquestions that will be addressed in one of the following chapters.

Chapter 2

Do goal-driven and stimulus-driven processes account for variation in context-based decision-making as measured by the *AX*-CPT? This chapter consists of two behavioral experiments with an adapted version of the *AX*-CPT. Both studies investigated the contribution of a goal-driven bias, induced by cue information and task instructions, and a stimulus-driven bias, induced by episodic bindings formed between cue and probe, to task performance.

Chapter 3

If both accounts explain variance in context-based decision-making, are these accounts also supported by the expected neural activation? This chapter contains an fMRI study

using an adapted version of the *AX*-CPT. Again, the contribution, and neural correlates, of a goal-driven vis-à-vis a stimulus-driven bias in task performance was studied.

Chapter 4

What is the modulating influence of a positive-affect induction on control processes in context-based decision-making? This chapter includes an ERP study with a classic *AX*-CPT, which addressed the effect of induced positive affect on proactive, reactive and evaluative control. The positive affect induction was assumed to increase DA levels in the brain.

Chapter 5

What is the modulatory effect of DA on reward-based decision-learning processes associated with specific striatal structures, i.e. caudate and putamen?

Chapter 6

What is the effect of subthalamic nucleus stimulation on reward-based decision-learning processes associated with caudate and putamen?

Chapter 5 and 6 tap more closely into the effect of BG modulations of reward-based decision-learning. Both chapters involved PD patient studies with medicated patients and patients treated with deep brain stimulation performing a reward-based decision-learning task.

The five empirical chapters are either published, under revision, or submitted in international psychological journals. They are inserted in this thesis in their original, submitted or published form. To acknowledge the contributions of several co-authors to these articles, a list of references is presented.

Chapter 2

Van Wouwe, N.C., Band, G.P.H., & Ridderinkhof, K.R. (2009). Proactive Control and Episodic Binding in Context Processing Effects. *Acta Psychologica, 131 (3),* 245-253.

Chapter 3

Van Wouwe, N.C., Band, G.P.H., Pannebakker, M.M., de Bruin, L.C., Ridderinkhof, K.R. & Hommel, B.. Goal-driven versus stimulus-driven components of biased decision-making: A functional neuroimaging study of the *AX*-CPT. *Manuscript submitted.*

Chapter 4

Van Wouwe, N.C., Band, G.P.H., & Ridderinkhof, K.R. (2009). Positive affect modulates flexibility and evaluative control. *Journal of Cognitive Neuroscience (in press).*

Chapter 5

Van Wouwe, N.C., Band, G.P.H., Wylie, S.A.W., van den Wildenberg, W. P. M., & Ridderinkhof, K.R. Learning processes associated with caudate and putamen are improved by dopaminergic modulation: evidence from Parkinson's disease. *Manuscript in preparation.*

Chapter 6

Van Wouwe, N.C., Wylie, S.A.W, van den Wildenberg, W. P. M., Band, G.P.H., Abisogun, A., Elias, W.J., Frysinger, R., & Ridderinkhof, K.R.. Deep brain stimulation of the subthalamic nucleus improves reward-based decision-learning in Parkinson's disease. *Manuscript in preparation.*