1 Introduction

Human behaviour is goal directed. Our ability to selectively adapt to the different situations and tasks of daily life in order to achieve these goals is crucial for efficient behaviour (Norman & Shallice, 1986). It is a typical human ability to perform actions according to intentions in an internally rather than externally- or stimulus-driven way. According to many researchers, action goals are believed to play an essential role in the acquisition, control, and planning of these intention-based actions (e.g., Hommel, 2003; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1997). Thus, some kind of cognitive control system is thought to protect the action process from disturbing environmental influences to achieve a target goal. Therefore, the system has to be stable to keep track of the intended goal but also flexible enough in order to be able to integrate actual environmental incidents, if required by the situation. For example, the smell of a fire should interrupt any actual goal and lead to another behaviour like escaping or putting out the fire. Goschke (2000, 2003) called this a *stability-flexibility dilemma* that has to be solved by cognitive control.

A good example to describe the requirements on cognitive control mechanisms are action slips in our everyday behaviour (Norman, 1981; Reason, 1979, 1990). The classical, and probably most cited, example is supplied by William James's anecdote of the man who goes to his bedroom with the goal of dressing for dinner and finds himself in bed with his pyjamas on. This example shows that external stimuli may trigger and bias our behaviour in a way that may lead to unintended actions. These so-called *capture errors* are an example of how external stimuli and internal goals compete for a particular behaviour. It is also a good starting point to examine the underlying mechanisms of a cognitive control system and the so-called executive functions.

1.1 Models of executive functions

Executive functions are considered to be the most powerful system in terms of human cognition and often compared to some kind of intelligent agent like a homunculus. It can be seen as a unified system with multiple functions or simply as an agglomeration of independent though interacting control processes. It includes a huge number of different neuropsychological and cognitive functions. In general, executive functions can be described as a system involved in mental and conscious planning, cognitive flexibility, abstract thinking, rule acquisition, initiating appropriate actions, inhibiting inappropriate actions and selecting relevant sensory information (Burgess & Shallice, 1996; Lezak, 1995; Shallice, 2004).

1.1.1 Multicomponent model of working memory

One of the first models in which executive functions are described is Baddeley & Hitch’s (1974) multicomponent model of working memory. Working memory can be seen as a system that is necessary for holding and manipulating information while performing a wide range of tasks including learning, reasoning and comprehending (Baddeley & Hitch, 1974). Baddeley and his colleagues proposed a modular working memory system including two limited capacity, short-term buffers or slave systems (Baddeley, 1986; Baddeley & Hitch, 1974).In addition, they proposed that a “central executive” is responsible for the selection of processes, the application of processing capabilities, the supervision of information integration and coordinating the two slave systems (see Figure 1). Furthermore, Baddeley (1986) proposed that the central executive is responsible for the attentional control of working memory (Baddeley & Della Sala, 1998) and for directing attention to relevant information, suppressing irrelevant information and inappropriate actions and coordinating cognitive processes when more than one task must be done at the same time.

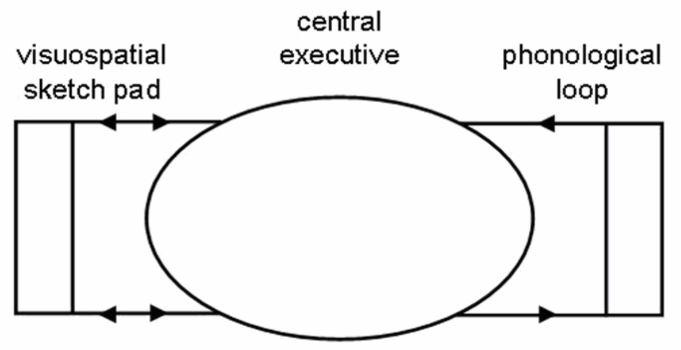

Figure 1: Baddeley & Hitch's model of working memory

One of the two limited capacity units is an auditory short-term buffer, referred to as a *phonological loop*, which contains speech-based representations (Baddeley, Lewis, & Vallar, 1984; Vallar & Baddeley, 1984). This stores phonological information and to prevent its decay by silently articulating its contents, thereby refreshing the information in a rehearsal loop. The other slave system is the *visuospatial sketchpad*, which is a limited capacity buffer that holds encoded spatial data (Frick, 1988; Morris, 1987). It can be used, for example, for constructing and manipulating visual images, and for the representation of mental maps. The sketch pad can be further divided into a visual subsystem that deals with, for instance, shape, colour and texture and a spatial subsystem that deals with location (Logie, 1995; Quinn & McConnell, 1996; Smith & Jonides, 1995) and probably a kinaesthetic component (Smyth & Pendleton, 1990).

Baddeley (2000) extended the model by adding a fourth component, the *episodic buffer*, which provides temporary storage of information held in a multimodal code, and is capable of binding information from the subsidiary systems, and long-term memory into a unitary episodic representation. This *episodic buffer* resembles Tulving's (1972) concept of episodic memory, but it differs in that the episodic buffer is just a temporary store. In general, the model is a combination of short term storage capacity and processing control. There was another advancement which was later integrated into the original model (Baddeley, 2000; Baddeley, Chincotta, & Adlam, 2001): the *phonological loop* and the *visuospatial sketchpad* should both have bidirectional connections to long-term memory. Thus, both have access to information stored in long-term memory and can also store information there.

In general, in this multicomponent model the initial specification of the central executive was a little bit vague. Baddeley (1996) described it as some kind of ragbag into which could be stuffed all the complex strategy selection, planning and retrieval checking, when a subject performs even a very simple task. The central executive is thought to act as an attentional controller, selecting certain streams of incoming information and rejecting others. The final executive capacity to be considered was the ability to select and manipulate information in long-term memory (Baddeley, 1996). Therefore, the model was extended to describe the central executive in more detail. The central executive in the extended model is based on Norman and Shallice's *Supervisory Attentional System* (SAS; Norman & Shallice, 1986) which in the meantime might be part of the probably most cited and most powerful models of attentional control.

1.1.2 Model of attentional control

The basic unit of this model is the schema (i.e. an action or thought schema). The word schema comes from the Greek word "σχήμα", which means *shape* or more generally *plan*. Executive functioning in the *model of attentional control* is based on a system thought to be involved in handling novel situations outside the domain of some of our "automatic" psychological processes that could be explained by the reproduction of learned schemas. These schemas represent special condition-action pairs. Norman & Shallice (1986) outlined various situations where routine activation of behaviour would not be sufficient for optimal performance which can be summarized as: (1) Situations that involve planning or decision making, error correction or troubleshooting, (2) situations where responses are not well-learned or contain novel sequences of actions, (3) dangerous or technically difficult situations, and (4) situations which require the overcoming of a strong habitual response or resisting temptation.

Under these circumstances, schemas have to control a specific overlearned action or skill. The difference to Baddeley's model of working memory is that the critical process of selection of schemas that are to be 'run' involves two stages that are distinct processes. The first stage involves *Contention Scheduling* (CS), which is an automatic (bottom-up) process, involved in both routine and non-routine selection. The CS-System has the function of ensuring, in a routine way, the efficient use of the limited effectors and cognitive resources, given that competition exists to use these resources for many different purposes. The CS-system selects one out of a restricted number of compatible schemas in a way that it can control cognitive resources until their goal is achieved, unless a much higher priority schema is triggered. This selection process occurs when activation in a schema control unit reaches threshold and is maintained by mutual inhibition between the other surrounding, incompatible units, even if perceptual or short-term memory triggers are no longer present. Additionally, schema-produced activation leads to an activation of its sub-schemas, which can be selected in advance if the appropriate trigger occurs.

If the CS-triggered routine operations are insufficient and lead to unexpected failures in behaviour (e.g. if there is a mismatch between behavioural feedback and the expected behavioural effect), the *Supervisory Attentional System* (SAS), which involves consciously controlled (top-down) non-routine selection of schemas, becomes active. In this case, the schema selection process becomes slower, but more flexible, instead of fast, routine and rigid. This process involves biasing the operations of CS by additional activation of appropriate schemas. For instance, if an activated schema has to be inhibited or if no link between external triggers and an appropriate schema exists. The SAS generates new schemas and in this way learning occurs if new schemas are integrated in the hierarchical network of the CS. Behavioural control therefore emerges from an interaction between these two subsystems. The SAS can also override the CS when necessary - for example, the ring of a telephone will cause priming of 'answer the phone' behaviour by the CS, but it might be appropriate for the SAS to override this if the telephone belongs to someone else (see Figure 2 on the next page). Thus, the model of Norman & Shallice differentiates between *bottom-up* exogenous triggers and *top-down* endogenous control. The SAS reflects intentional control rather than reflexive responses to stimuli.

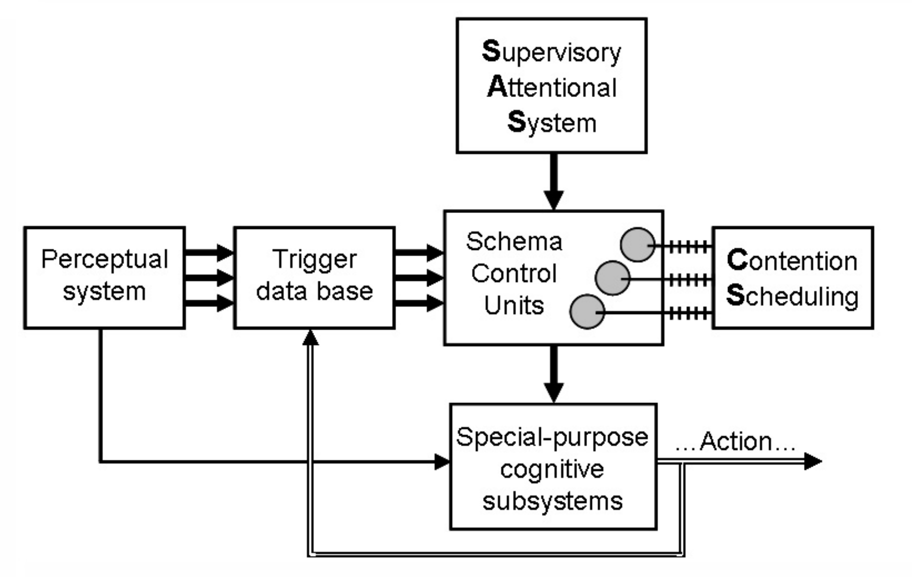

Figure 2: A simplified version of Norman & Shallice's (1986) model of attentional control representing the flow of information. The solid arrows represent activating input; crossed lines represent the primarily mutually inhibitory function of contention scheduling. The term 'special-purpose cognitive subsystems' refers to specific units involved in schema operations for both, action and thought schemas. In the latter case schema operation involves placing information in short-term stores that can trigger data bases.

However, the same criticism of Baddeley's central executive holds for the SAS: both models propose something like a clever homunculus that decides which schemas will be activated or inhibited. Neither of these models makes clear assumptions on how this decision process may take place. However, they show some basic properties that have been integrated into a series of other theories, i.e. action schemas that get automatically activated. We can find such *if-then rules* in production-system architectures (Anderson, 1976; Laird, Newell, & Rosenbloom, 1987; Newell, 1973) and as a kind of associations in connectionism models (Grossberg, 1988; Rumelhart & McClelland, 1986). Also, principles of control, i.e. the interaction between *bottom-up* and *top-down* processes, can be found in other cognitive models, for example, to describe the Stroop effect (Cohen, Dunbar, & McClelland, 1990) or in the stochastic race model of visual attention (Bundesen, 1990). Current models of cognitive control describe this construct as a result of learning, attention, and motivational processes (Duncan & Miller, 2002; Goschke, 2003; Mayr, Awh, & Laurey, 2003; Miller & Cohen, 2001).

1.1.3 Guided activation theory

Norman & Shallice’s SAS model based on studies with patients that suffer frontal brain lesions; therefore it describes the functions of the frontal lobe in an indirect way only. In contrast, Miller and Cohen’s “guided activation theory” (Miller & Cohen, 2001) focuses more directly on brain functions itself, especially those of the prefrontal lobe.

The guided activation theory proposes that the prefrontal cortex (PFC) stores representations of task-specific rules, attentional templates and goals. It ‘directs’ the activation of goal-related representations that are stored in the posterior cortex. This ‘guided activation’ of posterior representations is important in learning new rules and behaviour. The repetitive activation of the same pathway creates stronger associations between involved areas. Thus, during this process the role of the PFC in guiding posterior representations lessens – it might get virtually nil with frequently used rules and behaviours. Miller and Cohen use the metaphor of a switch operator, determining which railway tracks a train will use to describe the role of the PFC. By this analogy, if a train always uses the same track, then the switch operator is no longer necessary. Thus, this theory is consistent with both the PFC’s function of behavioural control and its connectivity with other brain regions (see below).

Specific assumptions can be derived from that theory: First of all, the PFC should be activated only in conjunction with the posterior cortex. Secondly, PFC involvement should increase as processing demands increase – which has been already shown in studies of cognitive control (Frith, Friston, Liddle, & Frackowiak, 1991; MacDonald, Cohen, Stenger, & Carter, 2000). However, this notion is inconsistent with the evidence of neural responses in the PFC to known actions (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti et al., 1996).

Miller and Cohen claim that the PFC is also important for the integration of information across stimulus domains (O'Scalaidhe, Wilson, & Goldman-Rakic, 1997; Rao, Rainer, & Miller, 1997; Wilson, O'Scalaidhe, & Goldman-Rakic, 1993). However, it is not clear yet how new representations generated in the PFC are transferred to the posterior regions for long term storage.

The models outlined above propose an “executive controller”. However, they fail to make clear how it works. Primarily, they cannot explain how intentions control our behaviour. To this end, cognitive psychologists developed a paradigm which should make it possible to examine executive functioning in healthy subjects by measuring their behavioural performance like reaction times and accuracy while intentionally controlling their behaviour – the task switching paradigm.

1.2 The Task-Switching paradigm: development and theories

The task switching paradigm provides an experimental framework for the systematic study of processes that are related to executive functions and cognitive control. It focuses on the ability to reconfigure mental resources while shifting from one task goal to another. The basic finding is that switching from one task to another is associated with a substantial increase in reaction time (RT) and a higher error rate than when repeating the same task (Allport, Styles, & Hsieh, 1994; Biederman, 1973; De Jong, 2000; Goschke, 2000; Jersild, 1927; Mayr & Keele, 2000; Meiran, 1996; Monsell, 1996; Rogers & Monsell, 1995; Spector & Biederman, 1976; Wylie & Allport, 2000). This difference in performance between switch and repeat trials is usually called switching costs or sometimes labeled as *task repetition benefits*.

There are a lot of different approaches to study executive control processes with the task switching paradigm and the next section should give an overview of the development of the task switching paradigm and the theories of the underlying mechanisms that cause switch costs.

1.2.1 The first Task-Switching experiments

Probably the first study on task switching is from Jersild (1927). He compared the processing time between heterogeneous itemlists, in which subjects alternated between different tasks, to the performance when processing homogeneous itemlists, in which subjects repeated just one task. The task in the homogeneous itemlists might be always to add 6 to a given digit, whereas in the heterogeneous itemlist subjects might have to alternate between adding 6 and then subtracting 3 from a given digit. Switch costs are represented by the difference in reaction times