Fractionation of executive function : a developmental approach

Huizinga, M.

Publication date
2006
Document Version
Final published version

Citation for published version (APA):
Huizinga, M. (2006). *Fractionation of executive function : a developmental approach*. [Thesis, fully internal, Universiteit van Amsterdam].

Fractionation of executive function:
A developmental approach

Fractionation of executive function:
A developmental approach

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op dinsdag 17 oktober 2006, te 12:00 uur

door
Mariëtte Huizinga
geboren te Amstelveen

Promotiecommissie:

Promotor: Prof. dr. M.W. van der Molen

Overige leden: Prof. dr. E.H.F. de Haan
                   Prof. dr. H.L.J. van der Maas
                   Prof. dr. K.R. Ridderinkhof
                   Prof. dr. B.A. Schmand
                   Prof. dr. M.C. Welsh

Faculteit der Maatschappij- en Gedragswetenschappen
Universiteit van Amsterdam
Afdeling Psychologie

Contents

1

2

Chapter 1   ●   General introduction

1.1.1   Executive function

Executive function (EF), or the ability to control thoughts and actions in order to achieve a future goal, is important in non-routinized situations. More specifically, EF is needed in carrying out a task that is complicated or novel, requiring sustained conscious attention and effort, planning and strategic thinking, the evaluation of feedback, and the flexible adjustment of behavior to rapidly changing demands of the environment (Miller & Cohen, 2001; Zelazo, Muller, Frye, & Marcovitch, 2003). In other words, EF is needed for example when one plans a series of events; or when one has to weigh alternatives and decide on a course of action; or when one has to flexibly change plans or actions; or when one has to perform a mental calculation; or when one has to evaluate an argument.

Importantly, improvement in EF is assumed to underlie development in a broad range of intellectual and social behaviors when children grow older (Case, 1985; Flavell, 1971; Siegler, 1983). For example, EF has been shown to be related to school performance, reading, mathematical and problem solving skills (Bayliss, Jarrold, Gunn, & Baddeley, 2003; Cowan, Saults, & Elliott, 2002; Swanson, 2006), and social-emotional development (Blair, 2003; Eslinger, Flaherty-Craig, & Benton, 2004; Hill & Frith, 2003).

1.1.2   A neuropsychological perspective on executive function

Research on EF has its historical roots in clinical neuropsychological investigations of patients with damage to the prefrontal cortex (PFC). One landmark case can be traced back to "4.30 PM on September 13, 1848, when an accidental explosion at a railroad construction site in Cavendish, Vermont hurled an iron tamping bar through the head of a 25-year-old foreman named Phineas Gage" (Harlow, 1848; in Mesulam, 2002, p. 8). Although Phineas Gage suffered extensive damage to the frontal lobes, he survived the accident, and intelligence, memory, speech, sensation, and movement all recovered. However, following the accident he changed from a responsible, socially well-adjusted person into a person with a disturbed personality, who experienced severe problems in the control and regulation of his behavior, which rendered him unable to organize future activity or hold gainful employment (Tranel, 2002).

The literature is replete with case studies similar to that of Phineas Gage. Neuropsychological research indicates that these patients with lesions to the PFC show normal performance on IQ tasks (e.g., Shallice & Burgess, 1991), but are impaired on complex EF tasks, such as the Wisconsin Card Sorting Task (WCST), and disk-transfer tasks, such as the Tower of Hanoi (but see Duncan, Burgess, & Emslie, 1995). The

WCST requires participants to infer, by trial and error with feedback, the relevant sorting rule out of three possible sorting rules (i.e., rules based on the color, shape, or number of the stimulus). After ten correct sorts, the sorting rule is changed without warning, thus requiring participants to identify the newly relevant sorting rule. A commonly used indicator of WCST performance is perseveration, which is defined as the persistence in responding according to the previous, but currently no longer relevant, sorting rule (Heaton, Chelune, Talley, Kay, & Curtis, 1993, p. 8). Performance on this task has shown to be sensitive to PFC damage, as PFC patients commit more perseveration errors compared to healthy individuals (e.g., Anderson, Damasio, Jones, & Tranel, 1991; Stuss et al., 2000). The ToH involves a series of different problems; the object of the task is to move four disks from a start state to a goal state in the fewest possible number of moves, while following a specified set of rules. These rules, which specify the permitted moves, are 1) only one disk can be moved at a time; 2) only the top disk on any peg can be moved; 3) only a smaller disk can be placed on a larger disk (Simon, 1975). Successful task performance requires participants to plan their moves to achieve the goal, and to inhibit the tendency to focus on short-term goals. Patients with lesions to the PFC have been reported to fail at recognizing and/or resolving ultimate goal–short-term goal conflicts (e.g., Goel & Grafman, 1995).

Superficially, children resemble adult PFC patients in the performance on the WCST. That is, on the WCST children also display perseverance and have difficulties with set-maintenance (e.g., Chelune & Baer, 1986; Levin et al., 1991; Welsh, Pennington, & Groisser, 1991). On the ToH, children show pronounced impairments in planning (e.g., Bull, Espy, & Senn, 2004; Welsh et al., 1991). The impairments in children's task performance are interpreted in terms of a relatively slow maturation of the PFC (Dempster, 1992; Stuss, 1992). Neurophysiological studies established protracted brain maturational processes within the PFC, enhancing neuronal transmission, such as synaptic pruning and myelination (Giedd et al., 1999; Gogtay et al., 2004; Huttenlocher, 1979; Sowell, Delis, & Stiles, 2001; Yakovlev & Lecours, 1967). The observed improvement of EF during development has been attributed to PFC activation becoming less diffuse and more focal (e.g., Amso & Casey, 2006; Casey, Tottenham, Liston, & Durston, 2005). Together, neuropsychological studies have shown that PFC damage results in impairment in EF, as indicated by performance on standard laboratory tasks, and that children's performance resembles that of PFC patients.

1.1.3   Fractionation of executive function

The interpretation of the "PFC hypothesis" of EF is complicated by the apparent dissimilarities in task performance of PFC patients, suggesting that the PFC involves a series of fractionated processes. For example, the performance of some PFC patients may

be impaired on the WCST, but normal on the ToH or vice versa. Also, impaired performance is not unique to PFC patients (for a review, see Stuss, 2006). Such inconsistencies have inspired researchers to examine the organization of EF. Several theories suggest EF to reflect a single, unitary mechanism, which does not include distinct sub-functions (Baddeley, 1986; Cohen & Servan-Schreiber, 1992; Kimberg, D'Esposito, & Farah, 1997; Norman & Shallice, 1986). In contrast, other lines of research provide evidence for the multi-faceted nature of EF, including distinct sub-functions with a focal neural correlate (Stuss, Shallice, Alexander, & Picton, 1995). For example, behavioral studies in a variety of samples, and using various standard EF tasks, have yielded low or nonsignificant correlations between tasks, while factor-analytic studies have tended to yield multiple factors (Brocki & Bohlin, 2004; Lehto, 1996; Lehto, Juujaervi, Kooistra, & Pulkkinen, 2003; Levin et al., 1996; Rabbitt, Lowe, & Shilling, 2001; Welsh et al., 1991). Miyake et al. (2000) studied the organization of EF, and its role in standard neuropsychological tasks. Using confirmatory factor analysis, they identified as distinct, but correlated factors three commonly postulated EF components Working Memory, Shifting and Response Inhibition. Importantly, Miyake et al. (2000) applied a latent variable approach, which facilitates the examination the organization of EF in terms of the variance that the tasks have in common, rather than in terms of isolated task performance. Also using a latent variable approach, Fisk & Sharp (2004) largely replicated the results of Miyake et al. (2000). Moreover, Miyake et al (2000) found that the EF component processes differentially predicted performance on the complex neuropsychological tasks: the latent factor Shifting predicted WCST performance, whereas the latent factor Inhibition predicted ToH performance.

The proposed fractionation of EF has been supported by neuroimaging studies that have provided evidence for the multi-faceted nature of EF, indicating that different regions within the PFC underlie different EF components. For example, the ability to maintain information in working memory has been found to recruit mostly lateral PFC (Narayanan et al., 2005; Smith & Jonides, 1999). In contrast, switching between tasks is thought to rely on medial PFC (Crone, Wendelken, Donohue, & Bunge, 2006; Rushworth, Walton, Kennerley, & Bannerman, 2004). Finally, the ability to inhibit responses was found to rely on orbitofrontal cortex (Aron, Robbins, & Poldrack, 2004; Roberts & Wallis, 2000). Thus, EF appears to be non-unitary, and distinct EF components are likely to contribute in different ways to complex task performance, on the WCST and the ToH.

1.1.4 Development of executive function

Developmental studies using standard neuropsychological tasks have shown that EF is characterized by a protracted course of development, beginning in early childhood and

continuing into adolescence (for reviews, see Diamond, 2002; Welsh, 2002). These studies have indicated that EF tasks are subject to distinct developmental trajectories, as reflected by the attainment of adult levels of performance at different ages. For example, studies that focused on the development of working memory capacity reported a gradual development throughout childhood, and into adolescence (e.g., Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Beveridge, Jarrold, & Pettit, 2002; Brocki & Bohlin, 2004; DeLuca et al., 2003; Gathercole, Pickering, Ambridge, & Wearing, 2004; Hitch, Halliday, Dodd, & Littler, 1989; Luciana, Conklin, Hooper, & Yarger, 2005; Luciana & Nelson, 1998; Luna, Garver, Urban, Lazar, & Sweeney, 2004). In addition, studies of the development of task switching abilities showed that the cost of switching between rules of behavior decreases until the age of about 12 (Cepeda, Kramer, & Gonzalez de Sather, 2001; Crone, Bunge, Van der Molen, & Ridderinkhof, 2006; Kray, Eber, & Lindenberger, 2004). Finally, inhibitory control was found to improve into late childhood (Bédard et al., 2002; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Durston et al., 2002; Klenberg, Korkman, & Lahti Nuuttila, 2001; Ridderinkhof, Van der Molen, Band, & Bashore, 1997; Van den Wildenberg & Van der Molen, 2004), or even into early adolescence (Williams, Ponesse, Schachar, Logan, & Tannock, 1999).

The interpretation of the developmental pathways is complicated by a number of factors. First, different tasks and subject samples are used across studies to measure the same EF component. Second, it is unclear whether the same EF construct is measured across age groups. Several researchers have demonstrated that we should focus on latent factors to understand the organization of complex constructs, such as EF (e.g., Bollen, 2002; MacCallum & Austin, 2000), as this strategy will facilitate the reliable assessment of differences in development across EF component processes.

The objective of this thesis is to explore the development of EF over the period from age 7 through young-adulthood. Understanding the nature of EF development provides a window on the normal functioning of the PFC, which will in turn contribute to a more comprehensive understanding of both the typical and atypical functional development of this highly interconnected brain region.

1.1.5 Outline of this thesis

The present thesis comprises five empirical studies that aim to understand the developmental pathway of EF from childhood to young-adulthood. Chapters 2 and 3 focus on the underlying processes of successful ToH and WCST performance. Chapter 2 concerns performance on the ToH in a sample of healthy young-adults. This study aimed to examine the contribution of recursive strategy knowledge to performance on the ToH. In addition, the influence of problem administration (items presented in ascending order of difficulty vs. items presented in random order) on task performance was examined. In

Chapter 3 we explored age-related change in performance on the WCST in terms of set-switching and set-maintenance processes in four age groups (i.e., 7-, 11-, 15-, and 21-year-olds). In addition, the contribution of the EF components Working Memory, Shifting and Inhibition to these processes was examined.

Most research into the development of EF component processes has focused on the development of working memory and inhibition. In contrast, the development of cognitive flexibility has received less attention. Therefore, Chapters 4 and 5 include several task-switching experiments in 7-, 11-, and 21-year-olds. In Chapter 4, the development of the ability to flexibly switch between two choice RT tasks was examined, as was the developmental trajectory of the ability to switch from stopping a response to executing a response. Chapter 5 presents the results of two experiments that aimed at explaining developmental changes in the ability to flexibly shift attention between two tasks in terms of the increased ability to resist interference from irrelevant information.

Finally, Chapter 6 involves a detailed examination of the organization and the development of the EF components Working Memory, Shifting and Inhibition, and their contribution to performance on the WCST and the ToL. We included multiple indicators of each EF component, such that we could study performance at the level of latent variables, according to a pre-specified model of EF. This chapter reports a multi-group confirmatory factor analysis including four large homogeneous age groups (i.e., 7-, 11-, 15-, and 21-year-olds).

All empirical chapters are published in, or at present submitted to, international journals. They have been inserted in this thesis in their original submitted or accepted form. To acknowledge the important contributions of the co-authors, a list of references is presented below:

Welsh, M. C., & Huizinga, M. (2005). Tower of Hanoi disk-transfer task: Influences of strategy knowledge and learning on performance. Learning and Individual Differences, 15(4), 283-298. (Chapter 2)

Huizinga, M., & Van der Molen, M. W. (2006). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. Manuscript submitted for publication. (Chapter 3)

Huizinga, M., & Van der Molen, M. W. (2006). Age-related change in switching from color to shape, and from stopping to going. Manuscript submitted for publication. (Chapter 4)

Huizinga, M., & Van der Molen, M. W. (2006). Age-related change in shifting attention between global and local levels of hierarchical stimuli. Manuscript submitted for publication. (Chapter 5)

Huizinga, M., Dolan, C. V., & Van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44, 2017-2036*. (Chapter 6)

Other papers related to the present thesis:

Alting von Geusau, N. A., Stalenhoef, P., Huizinga, M., Snel, J., & Ridderinkhof, K. R. (2004). Impaired executive function in male MDMA ("ecstasy") users. *Psychopharmacology, 175(3)*, 331-341.

Borst, J., Huizinga, M., & Van Rijn, H.D. (2006). Modeling intra- and interindividual differences in executive function tasks. Manuscript in preparation.

Hagenzieker, F., Huizinga, M., & Van der Molen, M.W. (2006). Sensation seeking behavior and executive function in MDMA users. Manuscript in preparation.

Huizinga, M., & Van der Molen, M. W. (2006). Developmental change in random number generation. Manuscript in preparation.

Welsh, M. C., & Huizinga, M. (2001). The development and preliminary validation of the Tower of Hanoi-Revised. *Assessment, 8(2)*, 167-176

Chapter 2 ● Tower of Hanoi disk-transfer task: Influences of strategy knowledge and learning on performance

Abstract

The Tower of Hanoi has become a popular tool in neuropsychology and cognitive psychology to assess a set of behaviors collectively referred to as executive function (EF). Substantial variability in performance on the Tower of Hanoi (ToH) disk-transfer task among normally functioning young adults, and potential contributions to these individual differences, were examined. In this expanded 60-problem version of the four-disk ToH, the degree to which problem administration (blocked vs. random) and strategy knowledge influenced overall performance and changes in accuracy across problems was examined. Eighty-seven college students were randomly assigned to a Blocked Group (problems given in ascending order of move-length) and a Random Group (problems given in a random order). After administration of the ToH task, participants described their problem solving and these verbal protocols were analyzed with regard to four elements of a strategic approach to problem solving. Problem administration order demonstrated no effect on task performance or on expressed strategy knowledge; however, strategy knowledge did predict performance on the ToH. An expected decrease in performance across trials was observed in the Blocked Group, and an increase in accuracy in the Random group indicated a learning effect. Strategy knowledge did not interact with these changes in performance across the items. These results suggest that external cues do not influence performance on the ToH to the same extent as individual differences in strategy induction relatively early in the problem solving process.

2.1 Introduction

The Tower of Hanoi (ToH) is a complex problem-solving task that has become a popular measure of the EF construct, which has been defined as "the ability to maintain an appropriate problem solving set for attainment of a future goal" (Welsh & Pennington, 1988; p. 200). This task has demonstrated sensitivity to prefrontal lobe function and dysfunction (e.g., Fuster, 1997; Glosser & Goodglass, 1990; Goel & Grafman, 1995; Goldstein & Green, 1995; Lezak, 1995; Stuss & Benson, 1986). However, the specific executive processes recruited for successful performance, and by implication impaired in prefrontal dysfunction, is the subject of current debate. Researchers have suggested that this task taps executive processes such as planning, working memory, and inhibition (e.g., Goel & Grafman, 1995; Roberts & Pennington, 1997), and empirical evidence supporting

these proposals has been converging in recent years. For example, research findings have suggested that inhibition (Welsh, Satterlee-Cartmell & Stine; 1999), working memory (Goel, Pullara & Grafman, 2001; Handley, Capon, Copp, & Harper, 2002; Numminen, Lehto & Ruoppila, 2001; Lock, Welsh, Adams & Kurtz, 2002; Welsh, Huizinga, Granrud, Cooney, Adams & Van der Molen, 2002), procedural learning (Bagley, Welsh, Retzlaff, Wolf & Bryan 2002; Davis & Klebe, 2001; Devine, Welsh, Retzlaff, Yoh & Adams 2001; Goldberg, Saint Cyr & Weinberger, 1990), and fluid intelligence (Devine et al., 2001; Lock et al., 2001; Numminen et al., 2001) each contribute to performance on the ToH task.

The ToH task has been termed a disk-transfer task, that is, in a series of different problems, the object is to move the disks from a start state to a goal state in the fewest number of moves as possible while following a specified set of rules. There are three rules that dictate moves that can and cannot be made: (1) only one disk may be moved at a time; (2) disks may be moved only to another peg (e.g., they cannot be placed onto the table or held in the hand while another disk is moved); and (3) a disk may never be placed on top of a disk smaller than itself. It has been assumed that these constraints force the participant to engage in planning activity in working memory, as well as the inhibition of intuitive, albeit maladaptive, moves (Pennington, 1994; Scholnick, Friedman & Wallner Allen, 1997). For example, to successfully solve the problem it is often necessary to temporarily block the desired ending arrangement in order to free up disks needed to achieve the goal pattern (Goel & Grafman, 1995). Successful task performance therefore requires that individuals plan their moves to achieve the goal state and inhibit the tendency to focus on short-term goals.

In his seminal analysis of the ToH task, Simon (1975) suggested that the constraints of this task are conducive to the spontaneous generation of several problem solving strategies that vary in effectiveness and may explain normal individual differences in performance. The optimal strategy is referred to as goal recursion and includes the following basic elements: (1) recognizing that the first subgoal is to move the largest disk to its goal position; (2) moving the smaller disks out of the way; (3) building a "sub-pyramid" stack of these smaller disks on the "open" peg; (4) moving the largest disk to it's goal position; and (5) repeating these steps with the "next-largest" disks and progressively smaller sub-pyramid stacks until the goal state is achieved. In it's most complex form, the goal recursion strategy involves the understanding that within the each major subgoal, i.e., building the subpyramid and delivering the largest disk to its final goal position, are smaller cycles of the five steps above. Simon (1975) proposed that the goal recursion strategy does not require perceptual updating with regard to the configuration of the disks on the peg, but simply the knowledge of where the person is in the stack of goals that must be achieved. In contrast, Simon (1975) suggested that the perceptual and sophisticated perceptual strategies include an appreciation of some subset

of the elements comprising the recursive strategy (e.g., recognizing the first subgoal), but the person must decide where to move the disks in order to sequentially deliver each largest disk to its goal peg. Importantly, when the person uses these perceptual strategies, there is no appreciation of building successively smaller pyramids on the "other" peg (neither source peg nor goal peg). In the "production statements" set forth by Simon to describe the sophisticated perceptual strategy, the movement of the smaller disks is perceptually-driven, rather than conceptually-driven, and often the result of an educated guess between two legal moves. In contrast, the goal recursion strategy provides a "plan" for how the smaller disks should be moved out of the way to clear the largest disk's path to its goal. This "plan" derives from the rule that a smaller subpyramid must be built on the open peg before the largest disk is moved to the goal peg.

One can also hypothesize that knowledge of the recursive rule guiding solution should reduce the demands on executive processes such as planning, working memory, and inhibition. Since Simon's pioneering work, there has been surprisingly little research on the problem solvers' knowledge of the recursive or perceptual strategy while engaged in the ToH task. There has been evidence that the recursive nature of the problem is one of the aspects that makes the task difficult for individuals to solve. For example, Welsh, Cicerello, Cuneo, and Brennan (1995) found a systematic pattern of errors and pause times that parallel each major recursive cycle within the move-path. Kotovsky, Hayes, and Simon (1985) found that, on an isomorph of the ToH task, participants tend to go through a large exploratory phase prior to "seeing" the full solution and then proceeding quickly to the goal. The verbal protocols of their participants suggested that they did not have the full solution in mind prior to the first move, and one can speculate that it was only during this exploratory phase that they discovered the recursive structure of the task. Goel et al. (2001) developed a computer model in which the ToH problems were solved via the perceptual strategy, and their results demonstrated the heavy working memory demands posed by longer-move problems. A main purpose of the current study was to contribute to our understanding of ToH solution by examining the general nature of strategy knowledge and the contribution of this knowledge to performance.

The search for the cognitive strategies underlying ToH performance is complicated by the fact that administration of this task varies considerably from laboratory to laboratory. Some variations include, among other things, the imposition of time limits, allowing either a single attempt or multiple attempts to correctly solve a single problem, and variation in orders of problem presentation. The approach to problem presentation has taken two basic forms. One form involves presenting only the original 3 or 4-disk, tower-to-tower to participants (e.g., Davis & Klebe, 2001; Goldberg et al., 1990; Squire, Cohen & Zouzounis, 1984; Welsh et al., 1995). The second form entails presenting a series of problems in order of ascending or descending order of difficulty (e.g., Borys, Spitz & Dorans, 1982; Welsh, 1991; Welsh et al., 1999; Welsh & Huizinga,

2001). In this case, as the participant progresses through the task, the problems increase or decrease in terms of the number of moves required to solve the problems successfully. Problems with a greater number of moves presumably place more demands upon working memory since participants are required to hold and manipulate more moves in memory before they overtly implement those moves with the physical ToH. It is important to note that the most complex ToH problem for a given number of disks (e.g., 15-move problem for four disks) can be decomposed into a series of shorter problems that are embedded within the larger problem, and it is some subset of these smaller problems that is typically presented in ascending order leading up to the most complex problem. The fact that the same errors continue to be repeated across these problems suggests that participants do not recognize the embedded structure of the ToH task as presented in this manner (Welsh, 1991). The current study was designed to examine the effect of experience with embedded problems on performance within the context of a 60-problem task that provides a substantial amount of this experience.

Given the variety of administration procedures characterizing the ToH tasks, it is critical to examine the influence of these procedures on performance and on the cognitive processes involved. The first goal of the present study was to explore whether the order in which the individual ToH problems are administered affects a) overall performance and b) performance at differing levels of difficulty. Specifically, a blocked administration, that is, in an ascending order of difficulty (e.g., 6-move problems, 7-move problems etc.), was compared to a random order of administration. Because the blocked order of administration capitalizes on the natural embedded structure of the ToH, it was predicted that participants in the blocked order group would exhibit better overall performance than would participants in the random order group.

A second goal of the study was to explore whether the two orders of problem administration affected the strategy knowledge expressed at the end of the large set of 60 ToH problems. Given that Simon (1975) identified the goal recursion strategy as the optimal method of problem solution on the ToH, would the order of problem presentation affect the participants' knowledge of the various aspects of this strategy? In a study by Schunn, Lovett and Reder (2001), the results indicated that explicit awareness of changing characteristics of their task was more predictive of appropriate adaptation of strategy than were individual differences in either working memory or inductive reasoning. External cues provided by administration order of the items is one method by which such self-awareness could be influenced. Therefore, we hypothesized that administering the problems in the blocked order would facilitate the discovery of an effective strategy because as problem difficulty (i.e., number of moves) increases, this necessitates another cycle or recursion through the strategy. We expected that participants in the random order group would have less explicit awareness of the elements of an

effective strategy (including aspects of goal recursion), given that the natural embedded structure is violated and the recursive nature of the task would be less obvious.

The problem-solvers' knowledge of an effective solution strategy may be conceptualized as an effect of the administration procedure experienced by the individual. Alternatively, one can view strategy knowledge as an individual difference variable that, in and of itself, contributes to performance on the ToH task. Performance on the ToH has been linked to inductive reasoning and other aspects of fluid intelligence (Devine et al, 2001; Lock et al, 2002; Welsh et al, 2002), as well as to the hypothetico-deductive reasoning characteristic of formal operational thought (Emick & Welsh, 2005). These cognitive processes may contribute to the likelihood of inducing a goal-oriented strategy and applying it effectively to the task, irrespective of problem order. Therefore, a third goal of the present study was to characterize participants in terms of their level of strategy knowledge, and to explore whether there were observable group differences in ToH performance.

Finally, the issue of learning across a large set of ToH items was of interest in light of the fact that this task has been referred to as a procedural learning task by some research groups (e.g., Davis & Klebe, 2001; Goldberg et al., 1990; Squire et al., 1984). Typically, those studies that utilize the ToH as a procedural learning task administer a single ToH problem repeatedly, and improvements in performance have been observed. However, in those studies administering multiple problems of increasing difficulty, a robust difficulty effect is typically found; that is, there is a substantial decline in performance across trials (e.g., Welsh, 1991; Welsh & Huizinga, 2001). This administration procedure clearly obscures the learning that may also be occurring across trials. A fourth goal of this study was to explore change in performance across 60 ToH trials: a decline in performance was expected in the blocked order group, whereas an increase in performance indicative of learning was expected in the random order group. It also was of interest whether strategy knowledge, as an individual difference variable, would influence the decrement in performance over trials in the blocked order group and/or the learning over trials in the random order group. If strategy knowledge is acquired over the course of the ToH trials, it is possible that the two strategy knowledge groups (i.e., high vs. low knowledge) will not differ in the early trials of the task, but progressively diverge in performance as the task continues.

To address these four issues, a sample of young adult volunteers were administered an expanded 60-item ToH tasks in one of two administrations: blocked order of random order. It was predicted that those participants in the blocked order group would exhibit better overall performance on the ToH, as well as a greater amount of strategy knowledge. In addition, it was expected that strategy knowledge itself, will relate to overall ToH performance and may influence the degree to which one observes changes in performance over trials.

2.2 Method
2.2.1 Participants

A total of 87 undergraduate students enrolled at a mid-sized university participated in the study. Upon volunteering for the study, the participants were questioned as to known diagnosis of any learning disabilities or head injuries. If the participant had a history of head injury and/or had a diagnosed learning disorder, he or she was excluded from the study. Six participants were excluded from the analyses for the following reasons: three participants (two females, one male) were excluded due to self-reports of past head injury, and three participants (three females) were dropped from the analyses due to extremely low scores (well below 2 standard deviations from the sample mean) and behaviors indicating a failure to understand the task. Therefore, the sample of participants included 81 students, 60 females and 21 males, with a mean age of 18.32 years (SD = .70). These students were compensated with course credit for their participation.

2.2.2 Apparatus

The Tower of Hanoi (Simon, 1975) consists of a flat board (40 x 15 x 2 cm) on which three vertical wooden pegs of equal diameter (1cm) and equal height (14.5 cm) are spaced equidistantly (12.5 cm). Four wooden disks of graduated size (13.5, 11, 8.5, and 6 cm diameter), each have one hole (1.3 cm in diameter) drilled through the center so that they fit onto any of the three pegs. A set of 17 x 22 cm cards in a three-ring binder displays the goal states of the individual items that are presented to the participant.

The ToH requires that an initial start configuration of disks across the three vertical pegs be transformed into a specific goal configuration of these objects, in the minimum number of moves. Disks must be moved according to a set of specified rules that constrain the manner in which these objects may be moved from peg to peg. These rules include the following: (a) only one disk may be moved at a time; (b) a disk may not be placed on the table or held in the hand while another disk is being moved; and (c) a larger disk may not be placed on top of a smaller disk.

2.2.3 Procedure

Participants were randomly assigned to either the blocked-order group or the random-order group by a coin toss. The task consisted of 30 tower-ending problems (e.g., all odd-numbered items on Figure 1; Welsh & Huizinga, 2001) and 30 flat-ending problems (e.g., all even-numbered items on Figure1; Welsh & Huizinga, 2001). The 60 items were all possible items derived from two possible different end-states (tower or flat), 10 different

move-lengths per item (6- through 15-moves), and three different starting pegs (Welsh & Huizinga, 2001). The blocked-order of administration involved presenting problems in an ascending order from 6-move problems through 15-move problems. The random-order of administration was developed by randomly selecting problems in an alternating sequence of tower-ending and flat-ending items. Participants were administered the ToH individually after an explanation of the three rules (see above) and two three-disk practice problems (one flat-ending and one tower-ending). Experimental items consisted of four-disk problems. For each item, the tester set up the start state on the ToH apparatus, and the participant was presented with a card that exhibited the goal state. Both this card and the tester indicated the number of moves required to achieve the goal state. Participants had to reach the goal in the designated number of moves on the first attempt, and there was no time limit imposed. Scoring involved awarding one point for each correct solution (i.e., transforming the start state to the goal state in the required number of moves).

At the end of the session, participants were questioned as to how they solved the problems. The experimenter set up a 4-disk, 15-move, tower-ending problem and asked the participant what should have been done to solve the problem. The experimenter then recorded the participant's response verbatim. Based on these written records of the participants' verbal protocol of their problem solving process, two trained, independent researchers assessed which, if any, of the four strategy components were articulated in the response given by the participant. The four components included: (a) move the largest disk to its goal first; (b) move the smaller disks out of the way; (c) build a "mini-tower" on the "open" peg; and (d) repeat the process to completion. The inter-rater agreement for the assessment of presence of these four elements was 95%, and all disagreements were resolved by mutual consensus prior to data analysis. A total recursive strategy score was calculated by assigning one point for each component present in the response.

2.3   Results
2.3.1   Task administration effect on performance

The prediction examined was that accuracy on the ToH task would be positively impacted by the blocked administration procedure in which 60 problems of 7- through 15-moves were presented in ascending order. In contrast, the random order of problem administration was expected to result in less accurate performance. The descriptive statistics for ToH performance in the two administration groups are presented in Table 1. Three independent variables were the focus of this analysis: Administration Group (blocked order vs. random order), Move Length (7 to 10 moves vs. 11 to 15 moves), and Goal State (problems with a tower-ending goal vs. problems with a flat-ending goal). A 2 (Administration Group) x 2 (Move Length) x 2 (Goal State) mixed model ANOVA identified significant main effects of Move Length, $F(1, 79) = 177.51$, $p < .05$, $MSE =$

867.11, and of Goal State, $F(1, 79) = 69.83$, $p < .05$, $MSE = 221.10$. There was no main effect for Administration Group, $F(1, 79) = 1.44$, $p > .05$, $MSE = 25.09$, nor did this factor interact with any other independent variable in the analysis. For both administration groups, the shorter move-length problems and the tower-ending problems were solved more accurately. The Move Length x Goal State interaction was significant, $F(1, 79) = 35.64$, $p < .05$, $MSE = 124.18$. As seen in Figure 1, there was a greater difference between the shorter move-length problems and the longer move-length problems for the flat-ending problems than for the tower-ending problems. Independent t-tests at each Move Length level confirm that there is no significant difference between goal states at the shorter move-length level ($t(80) = -1.76$, $p < .05$); whereas, there was a significant difference between goal states at the longer move-length level ($t(80) = -8.62$, $p < .05$).

	ToH total score overall	ToH total score on Tower-ending items	Total ToH score on Flat-ending items	ToH score on 7- to 10-move items	ToH score on 11- to 15-move items
Blocked	44.3 (8.4)	24.1 (4.3)	20.2 (4.8)	25.0 (3.6)	19.3 (5.7)
Random	42.1 (8.3)	22.3 (4.5)	19.7 (4.6)	24.8 (3.5)	17.3 (5.7)

Table 1. Means and SDs (between parentheses) of ToH scores per Administration Group

2.3.2   Analysis of strategy knowledge

The following analyses were conducted to examine the nature of the strategy knowledge expressed by the participants in the post-task interview, as well as the effect of the blocked vs. random order of problem administration on the amount of strategy knowledge expressed. The a priori prediction was that the blocked order of administration would result in greater strategy knowledge. In addition, strategy knowledge was analyzed as an individual difference variable to explore the differences in ToH performance between participants with relatively greater strategy knowledge vs. those with relatively less knowledge.

A 2 (Administration Group) x 4 (Strategy Element) mixed model ANOVA was conducted to examine differences in knowledge for the two administration groups, and for the four types of strategy elements: (1) move the largest disk to its goal first; (2) move the smaller disks out of the way; (3) build a "mini-tower" on the "open" peg; and (4) repeat the process to completion. The main effect of Administration Group was
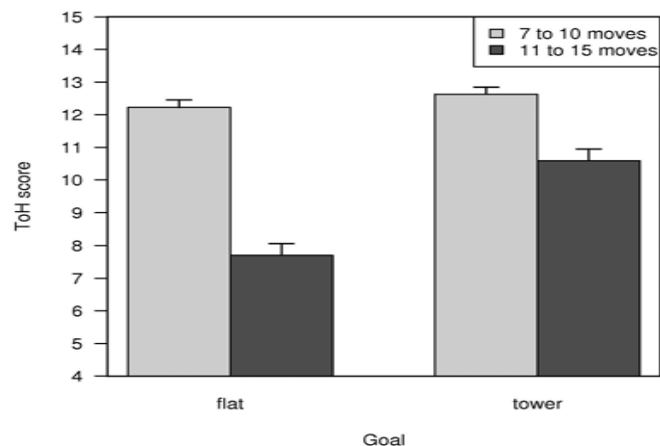
Figure 1. Mean ToH score (number of correctly solved items) and standard errors as function of Goal State and Move Length

nonsignificant, $F (1, 79) = .24$, $p > .05$, $MSE = .09$, and Administration Group did not significantly interact with the Strategy Element variable. Therefore, there was no effect of task administration type, random vs. blocked, on overall recursive strategy knowledge or on knowledge of the specific components of the strategy. There was a significant main effect of Strategy Element, $F (3, 237) = 15.83$, $p < .05$ $MSE = 2.38,$, indicating that there were differences in the degree to which the four elements were expressed in the verbal protocols of the participants. Given that the mean of the dichotomous score (1 or 0) for each strategy element reflects the percentage of participants expressing that element in their protocols, each of the first two elements were referred to by 80% of the participants, the fourth element was mentioned by 62%, and the third element was described by only 45% of the participants. Within subjects contrasts indicated that response rate to the first two elements did not differ significantly; however, each differed significantly from element three and four, which also were significantly different from each other.

In order to explore the association between strategy knowledge and ToH performance, each participant was assigned to one of two groups, based on the mean strategy knowledge score ($M = 2.5$ points of 4 possible points) for the entire sample (i.e., collapsed across administration group). Participants with a strategy score of 2.5 and above were assigned to the High Strategy group ($N = 35$) and participants with a score less than 2.5 points were assigned to the Low Strategy group ($N = 46$). A 2 (Strategy

Group) x 2 (Move Length) x 2 (Goal State) mixed model ANOVA examined the differences between the two strategy knowledge groups on overall ToH performance, as well as the degree to which these differences might be specific to ToH problems of different move-lengths or goal state configurations. The main effect of Strategy Group was significant, $F (1, 79) = 10.27$, $p < .05$, $MSE = 161.13$, reflecting a higher overall ToH score in the High Strategy group ($M = 45.61$, $SD = 6.97$) than in the Low Strategy group ($M = 39.91$, $SD = 9.03$). Strategy group assignment did not interact with the type of ToH problem, with regard to move-length or goal state. Consistent with the main effect of Strategy Group, strategy knowledge significantly correlated with ToH accuracy, $r (80) = .45$, $p < .05$, when knowledge was analyzed as a continuous variable (i.e., recursive strategy score from 0 to 4).

In light of the different rates at which participants in the post-test interview expressed the four strategic elements, it was of interest to examine whether these elements also differed with regard to prediction of the total ToH score. Point biserial correlations were calculated between scores (1 or 0) on each of the four recursive strategy elements and the total ToH score, as well as among the elements themselves. Significant correlations were found between ToH total score and scores for Point One ($r (79) = .29$, $p < .05$: "move the largest disk to its goal first"), Point Three ($r (79) = .41$, $p < .05$: "build a "mini-tower" on the "open" peg"), and Point Four ($r (79) = .37$, $p < .05$: "repeat the process to completion"). Moreover, intercorrelations among the scores on the elements of recursive strategy were found, most notably an association between the score on Point Three and Point Four ($r (79) = .50$, $p < .05$). To identify which of these four points were most predictive of ToH performance, a stepwise multiple regression was conducted and the only variable to enter the equation was Point Three (standardized beta = .41). The model was significant, $F (1, 79) = 15.80$, $p < .05$, $R^2 = .17$.

3.3 Changes in performance over trials

The 60 trials of this expanded version of the ToH were divided into quartiles of 15 problems each and these four sections included different problems for the two administration groups. For the Blocked Group, the first quartile included 6- to 8-move problems, the second quartile included 8- to 10-move problems, the third quartile included 11- to 13-move problems and the fourth quartile included 13- to 15-move problems. For the Random Group, the average move-length for the problems in each quartile were 9.8 moves (first quartile), 11.2 moves (second quartile), 10.3 moves (third quartile), and 9.9 moves (fourth quartile). The a priori prediction was that performance would decline in the Blocked Group as the move-length of problems increased, and that performance would increase in the Random Group if learning was occurring across trials.

The degree to which administration group and time across task (i.e., the quartiles) affected performance was analyzed by means of a 2 (Administration Group) x 4 (Time) Mixed Model ANOVA. There was a significant effect of Time, F (3,237) = 8.291, $p <$ .05, MSE = 30.899; however, more importantly the predicted Time x Administration Group interaction was significant, F (3, 237) = 56.61, $p <$ .05 MSE = 210.97. As Figure 2 illustrates, the performance of the Blocked Group indicated an increase from quartile one to two, and then the expected decrease in the following two quartiles. Within-subjects contrasts demonstrated that there was a significant change in performance within each pair of adjacent quartiles (e.g., between quartiles 1 and 2, 2 and 3, etc.). In contrast, performance of the Random Group exhibited an increase in accuracy across problems in the Random Group. Comparing adjacent quartiles, within-subjects contrasts found that the only significant difference in performance was between quartile 2 and 3 (F (1, 38) = 30.37, $p <$ .01).
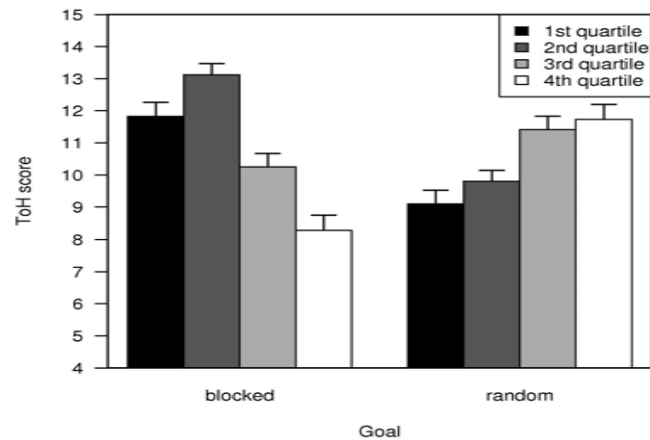

Figure 2. Mean ToH score (number of correctly solved items) and standard errors as a function of Administration Group and Time across the task

To examine whether strategy knowledge influenced the change in performance over trials, each of the administration order groups were analyzed separately. For the Blocked Group only, a 2 (Strategy Group) x 4 (Time) Mixed Model ANOVA identified main effects of Time (F (3, 114) = 49.85, $p <$ .05, MSE = 180.54) and of Strategy Knowledge (F (1, 38) = 4.88, $p <$ .05, MSE = 75.14). There was no interaction between

the Time and Strategy factors; thus, both the high and low strategy groups demonstrated a similar decrease in accuracy across the quartiles and the High Strategy Group maintained superiority in performance over the Low Strategy Group at each quartile. For the Random Group only, a 2 (Strategy Group) x 4 (Time) Mixed Model ANOVA identified main effects of Time (F (3, 117) = 17.31, $p <$ .05, MSE = 64.94) and of Strategy Knowledge (F (1, 39) = 5.43, $p <$ .05, MSE = 84.02). Again, there was no interaction between the Time and Strategy variables. As was the case for the Blocked Group, both the high and low strategy groups exhibited the same type of learning curve across the quartiles with the High Strategy Group demonstrating a higher accuracy score at each quartile. These performance trends over time can be seen on Figure 3.
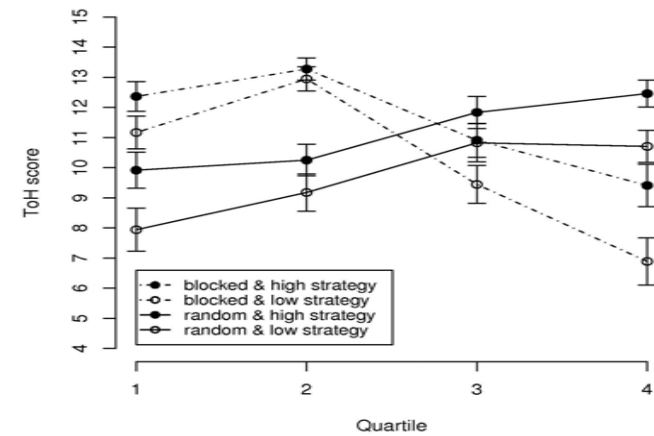
Figure 3. Mean ToH score (number of correctly solved items) and standard errors as function of Goal State and Move Length in two administration groups and two strategy knowledge levels as a function of Time across the task

2.4   Discussion

The purpose of this study was to gain a greater understanding of the mechanisms underlying performance on a structured problem-solving task, the Tower of Hanoi (ToH). By manipulating the order in which the sixty problems was administered, as well as examining the nature of the participants' strategy knowledge, the degree to which an appreciation of the inherent embedded structure of the ToH problem influences problem

solving success was explored. Taken together, the results suggest that the putative external cues provided for the participant through the order of problem administration does not aid in the recognition of this problem structure. Instead, it appears that strategy knowledge is an important individual difference variable that contributes to problem solving performance relatively early in the task.

It was predicted that the Blocked Order of administration would be related to greater problem solving success and greater recursive strategy knowledge because the recursive structure of the ToH problems would be more salient to the participants. Given that the earlier problems in the blocked order are simply smaller problems embedded within the larger 15-move tower-to-tower sequence, the participant can see the recursive strategy as it continues to "build" cycle by cycle. In contrast, the Random Order completely breaks down this natural embedded structure, presenting the problems in a completely random fashion. If participants are sensitive to cues provided by the embedded structure and use these to induce an effective strategy, then performance and/or strategy knowledge should have been significantly better in the Blocked Order group. This was not the case, as the Blocked Group participants did no better in overall performance, performance on longer-move or flat-ending goal problems, or in expressed strategy knowledge. As was found in previous studies in which the ToH problems were administered in ascending order, neither children (Welsh, 1991) nor adults (Welsh et al., 1995) appear to be sensitive to the embedded structure of the task. In both administration modes, strategy knowledge appeared to be acquired through extended "transactions" with the task, similar to that observed by Best (1990) in a different problem solving context.

With regard to strategy knowledge, it also was not affected by the manipulated variable of problem order. However, there was clear variability in the degree to which the four elements of the goal-oriented strategy were expressed in the post-test interviews, such that a "high strategy knowledge" and "low strategy knowledge" group could be identified. The high knowledge group scored significantly better on the ToH than the low knowledge group and degree of knowledge was positively correlated with the total score. Given that this was an individual difference variable, and was based on information gathered *after* the 60-problem task was over, one cannot conclude that greater strategy knowledge per se *caused* the superior performance. For example, it is entirely possible that one group of participants was more prone to inducing the strategy over the course of the task and the strategy was only fully realized toward the end of the large set of problems, right before the interview session. The analysis of change in performance over the four quartiles of the task provided a window through which one could examine this possibility. If indeed the strategy was being induced by some participants gradually across the 60-problem task, one might expect relatively equal performances of the low and high strategy groups on the first quartile (in both the Blocked and Random Orders) with divergence between the two groups in the later quartiles. That is, the increasing

---

strategy knowledge would be reflected in relatively less decrement with more difficult problems in the Blocked Order or relatively greater learning in the Random Order. Such an interaction was not revealed by the analyses, and instead, the high strategy group showed superior performance compared to the low strategy group at every quartile of the task. Thus, something about this high strategy group, greater knowledge or perhaps another factor such as self-monitoring (Schunn, Lovett, & Reder, 2001), was related to better performance in the first fifteen problems of the task. Given the novelty of the ToH task, it is unlikely that these participants possessed knowledge of this strategy prior to the administration of these problems; however, it is reasonable to hypothesize that these participants were more effective at inducing the rule-governed strategy very early in the problem solving process.

What cognitive processes might underlie this ability to quickly discover a successful strategy for solving a novel problem-solving task such as the ToH? Research in our lab has identified associations between performance on a 22-problem version of this task, the Tower of Hanoi-Revised (ToH-R), and measures of fluid intelligence, particularly inductive reasoning (Devine et al., 2001; Lock et al., 2002; Welsh et al., 2002). Additionally, the related process of formal operational thinking may facilitate the generation of effective problem solving strategies, and indeed, performance on the ToH-R has been found to correlate with formal operational reasoning (Emick & Welsh, 2005). Across several studies utilizing the ToH-R, the individual differences in performance in normal college-student samples are extremely consistent and undeniable (e.g., Welsh & Huizinga, 2001). It appears that a subset of young adults is more likely to induce a goal-oriented strategy, and this occurs over the course of the first dozen or so problems. One might speculate that there are more basic cognitive mechanisms that contribute to the ability to induce a rule-based strategy from patterns observed across these problems, and two candidate cognitive processes are working memory and inhibition. Patterns across problems are recognized only if one integrates information across time (Hambrick & Engle, 2003), and this presumably occurs in working memory. Various researchers have identified the important contributions of memory processes, such as working memory (Goel et al., 2001), and activation and priming (Altmann & Trafton, 2002) to performance on different versions of the ToH task. By "lesioning" the working memory representations of relevant subgoals, the researchers could simulate impaired performance on the longer-move problems, such as that observed in frontal damaged adults (Altman & Trafton, 2002, Goel et al., 2001). Similarly, inhibitory processes must be engaged by the person so as to resist being pulled in the direction of reasonable disk moves that are nevertheless inconsistent with the recursive strategy, and it is this demand that Goel and Grafman suggested was central to performance in an earlier study (Goel & Grafman, 1995). At this point, there is some controversy in the empirical literature regarding the contribution of working memory and inhibitory processes to ToH

performance. Whereas, a structural equation model proposed by Miyake, Friedman, Emerson, Witzki, and Howerter (2000) identified the inhibition factor as the best predictor of ToH performance, a structural model developed by Welsh et al. (2002) found working memory to be the stronger predictor. The differences in the measures of working memory, inhibition, and most importantly, the ToH, may contribute to the explanation for these incongruent results.

In addition to these explicit cognitive processes as underlying the induction of the strategy knowledge and performance on the ToH, many researchers have pointed to implicit learning as a factor in improved performance on the task. Typically this procedural learning is observed by continuous improvement in performance over repeated administrations of the same ToH problem. The classic study described by Cohen and colleagues (Cohen, 1984; Squire et al., 1984) demonstrated that amnesic patients exhibited this improved performance over repeated presentations, with no explicit memory of having done the task previously. Such evidence bolstered the claim that the ToH could be learned through implicit processes. Interestingly, in studies that employ the ToH as a procedural learning task, optimal performance on each problem (i.e., solving a 7-move problem in seven moves) is not required of the participants. In early administrations of the problem, a person might eventually solve a 15-move item in over 50 moves and procedural learning is indicated by the steady decline in number of moves, though perfect performance may never be achieved. Therefore, this type of ToH paradigm may not be appropriate for examining the discovery of an explicit strategy, such as goal recursion (Simon, 1975), which should lead to perfect performance. In our own lab we have found significant correlations between performance on a procedural learning task, Mirror Tracing (Gabrieli, 1998), and ToH-R performance (Bagley et al., 2002), with one study indicating somewhat equal contributions of both explicit learning and implicit learning to the explanation of variance in ToH-R scores (Devine et al., 2001). When a form of the recursive strategy was taught explicitly to participants before beginning the ToH-R task, these participants performed significantly better than a control group and, as expected, expressed greater knowledge of the recursive strategy (Lock et al., 2002). However, the "teaching group" performed no better than participants who were allowed to explore the task on their own without explicit strategy teaching, even though the "exploration group" exhibited less strategy knowledge. Equal performance, albeit with unequal explicit strategy knowledge, does suggest that at least a part of the performance on the ToH-R is related to implicit learning processes. The degree to which implicit versus explicit learning processes are involved in performance on various versions of the ToH continues to be a matter of debate. For example, Winter, Broman, Rose and Reber (2001) found evidence for problem solving on the ToH via declarative (i.e., explicit) strategies in the two cases of amnesia they studied, and concluded that the

ToH may not be a pure measure of implicit learning, as has been suggested in the literature.

It is of interest to consider not only how participants may induce move patterns and generate strategy knowledge, but also what types of strategy knowledge is induced. Of the four identified components of the strategy, the average number expressed in the post-test interviews was between two and three, and only 32% of participants stated all four elements. Those participants who understood the complete strategy had mean scores that were about three points higher on all performance measures. Interestingly, the 44% sample who expressed element three of the strategy exhibited scores that were indistinguishable from the group that recognized all four of the elements. The multiple regression analysis supports the conclusion that this component (i.e., building the smaller tower of disks on the "other peg" as the largest disk not currently on the goal peg is moved to the goal), expressed by the smallest number of participants, was nevertheless the most crucial element for success. Most participants appeared to understand that each major subgoal involved delivering the current largest disk to its goal as soon as possible, and to accomplish this, the smaller disks must be "moved out of the way". It was the notion that all smaller disks should be stacked in a pyramid configuration and where these disks should be stacked that posed the problem for most participants, and it is this lack of understanding that typically leads to the common errors one observes on the ToH task (Welsh, 1991; Welsh et al., 1995). Importantly, it is the understanding of the concept of "moving a smaller pyramid to the other peg" that distinguishes the recursive strategy from the sophisticated perceptual strategy (Simon, 1975) and, therefore, the best performers in this study appear to have a more fully realized recursive strategy. Both Altmann and Trafton (2002) and Welsh et al. (1995) have found the most frequent errors, as well as increases in pause times to occur at the beginning of each major recursive cycle, immediately after the largest disk has arrived at its goal destination. Such errors could be avoided by understanding the concept of the building a smaller pyramid characteristic, or, as has been suggested by other researchers, by an ability to "look ahead" several moves (Borys et al., 1982). The "look ahead" approach implies perception of the configuration of disks, which is closer to the sophisticated perceptual strategy as described by Simon (1975). Researchers disagree as to whether the task requires planning in the form of "look ahead," also known as "depth of search" or, alternatively, the ability to inhibit moves directly to the goal in order to make the counter-intuitive move (Goel & Grafman, 1995). In fact, it may be a combination of anticipating the results of future moves in working memory and inhibiting direct, but maladaptive, moves that allows the person to gradually appreciate the structure of the task and discover a systematic and effective strategy, such as goal recursion.

A limitation of this study is the fact that it is difficult to determine the precise nature of the covert strategy guiding the participants' problem solving, and the degree to

which the participants who expressed knowledge of all four elements fully realized the goal recursion strategy. Goel et al. (2001) state that the Simon's goal recursion strategy can be applied only in the classic tower-to-tower configuration of the ToH problem and not in intermediate states of the move path. This is certainly the case if one views the recursive strategy as a fixed algorithm, such as that programmed into computer models of ToH solution, such as ACT-R (Anderson & Lebiere, 1998). In our view, the basic elements of the recursive strategy (e.g., get largest disk to final goal, stack smaller subpyramid of disks on other peg, etc) can be applied successfully to intermediate problems, and even to flat-to-flat configurations, such as those included in the 60-item task. Given that perceptual information regarding the current and final goal configurations are referenced and monitored by the person, in Simon's framework our participants might be more appropriately classified as using the sophisticated perceptual strategy. This may be particularly true in the case of the more difficult flat-to-flat type of problem. It is true that the recursive strategy is optimally suited to solving a tower-ending problem, in which the largest disks are moved successively to a single goal peg. However, the same procedures can solve a flat-ending problem, which is an intermediate state on the way to the final tower-ending configuration. It is precisely the ability to "see" the flat-ending goal as an intermediate state of, or embedded in, the larger problem, that poses a serious difficulty for most participants. Anecdotally, participants who appear to have an understanding of the recursive strategy often behave as if they are caught by surprise when the current flat-ending goal is achieved a bit earlier than they expected. That is, they seem to be focused on the ultimate end of the goal recursion strategy: a stacked tower of disks on one peg. They begin executing the strategy and its rules in the deductive manner that Simon describes; that is, without constant reference to the current configuration of disks. Therefore, those participants who appear to have a fully realized goal recursion strategy may err by continuing past the intermediate goal on to the final tower configuration, unless they allow perceptual information (i.e., the intermediate goal on the card in front of them) to influence their responses. The current findings demonstrate that these flat-ending goal problems are clearly more difficult for these participants, perhaps because the goal recursion strategy that works so well in the tower-ending case does not apply as transparently in the flat-ending case.

It is not argued here that our participants have a full realization of the goal recursion algorithm as might be represented in a computer model. Instead, it is suggested that participants use the basic elements of the sophisticated perceptual strategy or the goal recursion strategy (i.e., if they have the concept of the "other peg") as a guide to achieve the subgoals in their appropriate order. For those using the sophisticated perceptual strategy, they may engage working memory process to "look ahead" in order to determine which would be considered the "other peg" in the more difficult flat-to-flat problems. A subset of participants was able to induce the rule of the "other peg" and then

deductively apply this rule to all of the problems. Those participants who went through this hypothetico-deductive process were presumably the ones to express their explicit understanding of the "mini-pyramid" and where to place it. Recall that it was this third element of the strategy that was most predictive of performance. Also consistent with this proposal is the conclusion drawn by both Anderson and Douglass (2001) and Altmann and Trafton (2002) that ToH performance does not follow the "perfect memory goal stack" mechanism, as is found in the their computer program known as the ACT-R cognitive architecture (Anderson & Lebiere, 1998). Instead, both research groups suggested that performance appears to depend on the priming, activation, and retrieval of subgoals from memory.

The general disk-transfer task paradigm that is reflected in the many variants of the ToH, and its distant cousin the Tower of London (Shallice, 1982), have become popular neuropsychological measures of EF presumed to be mediated by the prefrontal cortex (e.g., Stuss & Benson, 1984). Although the two types of tasks are often treated as isomorphic due to commonalities in their general structure (i.e., a start state and a sequence of moves to achieve the goal state), the TOL does not involve the inherent recursive structure that characterizes the ToH task. Therefore, it has been suggested that the ToH may tap inductive reasoning and the TOL may tap a more pure form of planning (Goel & Grafman, 1995); a hypothesis that was recently confirmed via structural equation modeling (Welsh et al., 2002). Before these tasks can be used effectively within a larger neuropsychological assessment battery, the construct validity of each must be examined further. Most importantly, the question of discriminant validity, or the differential EF that may be tapped by each task, should be explored (e.g., Welsh et al., 1999). The current study indicates that inductive reasoning as it applies to discovering the recursive strategy is an important cognitive mediator to success on an expanded version of the ToH and contributes to the individual differences in performance observed among normal, young adults. These results are convergent with the notion that there is significant overlap between frontally-mediated EF, such as working memory and inhibition, and the cognitive construct of fluid intelligence (Pennington, Bennetto, McAleer & Roberts, 1996). This perspective may substantially influence the development of future measures of EF in the field of neuropsychological assessment.

Chapter 3 ● Age–group differences in set–switching and set–maintenance on the Wisconsin Card Sorting Task

Abstract

This study examined developmental change in set-switching and set-maintenance on the Wisconsin Card Sorting Task (WCST), and sought to determine how executive function (EF) components (i.e., Working Memory, Shifting and Inhibition) may contribute to the observed changes on WCST performance. To this end, performance in four age groups (7-year-olds, 11-year-olds, 15-year-olds, and 21-year-olds) was measured on the WCST, and on three EF tasks assumed to tap Working Memory, Shifting, and Inhibition. The results showed that adult levels of performance were reached in 11-year-olds for set-switching, and in 15-year-olds for set-maintenance. A subsequent principal component analysis revealed that set-switching and set-maintenance loaded on two factors for 7-year-olds, but a single factor in the other age groups. Finally, regression analyses yielded a complex pattern of results concerning the prediction of set-switching and set-maintenance by the performance on tasks used to assess the EF components. The results were interpreted to suggest distinct developmental trends in set-switching and set-maintenance abilities required by the WCST.

3.1 Introduction

Executive function (EF) refers to a range of cognitive processes that subserve goal-directed behavior (Miller & Cohen, 2001; see also Luria, 1966; Shallice, 1982). Intact executive functioning is indispensable in novel or demanding situations that require the ability to control thoughts and actions (Stuss, 1992). Thus, EF underlies the ability to adjust behavior rapidly and flexibly to the varying demands of the environment (Zelazo, Muller, Frye, & Marcovitch, 2003). In the course of development, EF becomes increasingly more efficient. Developmental theories interpret this increase as an important manifestation of cognitive and emotional development (for reviews see Diamond, 2002; Welsh, 2002).

Probably the most frequently used experimental task to assess EF is the Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtis, 1993). The WCST requires participants to infer, by trial and error with feedback, a relevant sorting rule out of three possible sorting rules (i.e., the color, shape, or number of the stimulus). After ten correct sorts, the sorting rule changes without warning, requiring participants to find the newly relevant sorting rule. A commonly used

indicator of WCST performance is perseveration, which is defined as the persistence in responding to a previous, but currently no longer relevant, sorting principle (Heaton et al., 1993, p. 8).

The WCST was devised originally to assess deficits in EF in patients with brain damage (Berg, 1948), and the current consensus is that WCST performance relates to the integrity of the prefrontal cortex (PFC) of the brain (e.g., Demakis, 2003; Heaton et al., 1993; Lezak, Howieson, Loring, Hannay, & Fischer, 2004; Stuss & Knight, 2002). Neuroimaging studies report activation of dorsolateral PFC structures in successful WCST performance (e.g., Demakis, 2003; Monchi, Petrides, Petre, Worsley, & Dagher, 2001; Stuss & Levine, 2002). In addition, patients with PFC damage show an increased level of perseveration errors compared to normal controls (Anderson, Damasio, Jones, & Tranel, 1991; Barceló & Knight, 2002; Milner, 1963; Nagahama, Okina, Suzuki, Nabatame, & Matsuda, 2005; Stuss & Levine, 2002; Stuss et al., 2000).

Superficially, children's behavior on the WCST resembles the performance of PFC patients. More specifically, children are seen to also perseverate (Chelune & Baer, 1986; Chelune & Thompson, 1987; Heaton et al., 1993; Huizinga, Dolan, & Van der Molen, 2006; Kirk & Kelly, 1986; Paniak, Miller, Murphy, Patterson, & Keizer, 1996; Welsh, Pennington, & Groisser, 1991). Developmental studies, including the WCST, have established that adult level of performance is reached between late childhood and adolescence (Chelune & Baer, 1986; Chelune & Thompson, 1987; Huizinga et al., 2006; Levin et al., 1991; Welsh et al., 1991). The slow development of EF has been related to the relatively slow maturation of PFC (e.g., Casey, Tottenham, Liston, & Durston, 2005; Diamond, 2002).

3.1.1 Decomposition of executive function

One major difficulty in interpreting results from complex neuropsychological tasks such as the WCST concerns the lack of insight into the exact abilities that are necessary for successful performance (e.g., Miyake et al., 2000; Stuss & Levine, 2002), i.e., various EF component processes may contribute to task performance. Using structural equation modeling (SEM), Miyake et al. (2000) showed that EF can be divided into three distinct components: Working Memory, Shifting, and Inhibition (see also Fisk & Sharp, 2004; for a developmental study using SEM: Huizinga et al., 2006) Moreover, specific EF components were found to explain a significant proportion of variance related to performance on specific complex neuropsychological tasks. That is, Shifting predicted WCST performance, whereas Inhibition predicted Tower of Hanoi performance, a test used to assess planning abilities.

Recent neuro-imaging studies confirmed that EF is not unitary by showing that subcomponents of EF rely on distinct regions of PFC. For example, lateral PFC is

implicated in Working Memory (Narayanan et al., 2005; Smith & Jonides, 1999), whereas medial PFC is involved in flexible switching between tasks, and in overriding a previously relevant stimulus-response association (Crone, Wendelken, Donohue, & Bunge, 2006; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Rushworth, Walton, Kennerley, & Bannerman, 2004). Finally, the ability to inhibit responses was found to rely on orbitofrontal cortex (e.g., Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Roberts & Wallis, 2000).

In addition, recent developmental studies reported distinct developmental trajectories of EF component processes (for reviews, see Diamond, 2002; Welsh, 2002). Several studies revealed that the achievement of adult levels of Working Memory capacity is the outcome of development that proceeds well into adolescence (e.g., Beveridge, Jarrold, & Pettit, 2002; Brocki & Bohlin, 2004; DeLuca et al., 2003; Gathercole, Pickering, Ambridge, & Wearing, 2004; Luciana, Conklin, Hooper, & Yarger, 2005; Luna, Garver, Urban, Lazar, & Sweeney, 2004). In addition, adult levels of task shifting performance were found to be attained around the age of 12 (Cepeda, Kramer, & Gonzalez de Sather, 2001; Crone, Bunge, Van der Molen, & Ridderinkhof, 2006; Huizinga & Van der Molen, 2005; Kray, Eber, & Lindenberger, 2004). Finally, adult-levels of inhibitory control were observed to be reached around the age of 12 (Bédard et al., 2002; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Durston et al., 2002; Ridderinkhof & Van der Molen, 1995; Van den Wildenberg & Van der Molen, 2004), or early adolescence (Williams, Ponesse, Schachar, Logan, & Tannock, 1999).

3.1.2   Set-switching and set-maintenance on the WCST

The requirements for successful performance on the WCST include (1) efficient switching to the new sorting rule on the basis of feedback (i.e., set-switching), and (2) retaining the current sorting rule in mind through varying stimulus conditions, while ignoring irrelevant aspects of the stimuli (i.e., set-maintenance; Barceló & Knight, 2002; Heaton et al., 1993). In a recent patient study, Barceló and Knight (2002) used an analogue of the WCST to examine the nature of impairments in WCST performance. They put forward an error scoring method that distinguished between set-switching and set-maintenance processes, which are active during the process of finding a (new) valid sorting principle. The set-switching process was indexed by two types of errors: (1) perseverative errors, which occur when a participant fails to switch to another sorting rule after receiving negative feedback on the previous trial (see also Heaton et al., 1993), and (2) efficient errors, which are related to the efficient testing of hypotheses during switching to a new sorting rule. Efficient errors occur when a participant switches to the wrong sorting rule in the second trial of an otherwise clear series requiring the new

sorting rule (i.e., series with no errors other than the first error indicating that the sorting principle changed; Barceló & Knight, 2002). The set-maintenance process was indexed by distraction errors, involving random failures to maintain set. Distraction errors occur when the sorting rule is missed continuously, or when there is only one isolated error in an otherwise clear series requiring the correct sorting rule (Barceló & Knight, 2002).

Earlier, Barceló (1999) reported data from an event-related-potential (ERP) study in normal adults, indicating that perseverative errors and distraction errors are associated with distinct networks in the PFC. Errors reflecting set-switching abilities were associated with the activation of a frontal-extrastriatal network, whereas set-maintenance abilities were associated with a frontal-central activation. Aside from the finding of increased perseveration during WCST performance in the PFC patients participating in their study, Barceló and Knight (2002) suggested that the observed decrease in the ability to maintain set contributed to the set-switching deficits shown by these patients. That is, these patients had difficulties in keeping their attention focused on the new correct sorting rule (in the presence of distracting stimulus features; see also Milner, 1963).

Recently, Crone, Ridderinkhof, Worm, Somsen, and Van der Molen (2004) obtained evidence suggesting the separability of set-switching and set-maintenance processes during WCST performance. They tested four age groups (8-9-year-olds, 11-12-year-olds, 13-15-year-olds, and young-adults) using an experimental analogue of the WCST, which involved spatially compatible and incompatible S-R mapping rules. This task requires the deduction of a correct sorting rule on the basis of feedback or a switch cue. Task performance was scored following Barceló and Knight (2002). Crone and colleagues observed distinct developmental trajectories for set-switching and set-maintenance abilities. More specifically, set-switching abilities developed during childhood and reached adult levels of performance at age 12, whereas set-maintenance abilities continued to develop into adolescence.

3.1.3 Present study

The question whether children use the same abilities as adults on the traditional WCST remains unanswered to date. The WCST is a complex task, on which adequate performance draws on multiple higher cognitive processes (Miyake et al., 2000). In addition, participants may adopt different learning strategies when searching for a correct sorting rule (e.g., Schmittmann, Visser, & Raijmakers, 2006). In the present study, we examined the development of WCST performance by integrating the approaches advanced by Barceló and Knight (2002) and by Miyake et al. (2000). Barceló and Knight (2002) conceptualized WCST performance in terms of two processes—set-switching and set-maintenance—that were found to rely on distinct areas in the frontal cortex. In the current study, we adopted the conceptualization advanced by Barceló and Knight (2002)

and scored WCST performance in terms of set-switching and set-maintenance abilities. Subsequently, we examined the relative contributions to these abilities of the three EF components, distinguished by Miyake et al. (2000), i.e., Working Memory, Shifting, and Inhibition. We tested children in three age groups (7-year-olds, 11-year-olds, 15-year-olds), and one group of young-adults (21-year-olds). All participants were tested using a computerized version of the standard WCST, and using three tasks assumed to tap the EF components Working Memory, Shifting, and Inhibition (see also Huizinga et al., 2006)

Working Memory was defined as the collection of cognitive processes that temporarily retain information in an accessible state, suitable for carrying out any mental task (Cowan, 1998). The essence of this component is the monitoring and coding of incoming information with respect to relevance, and the replacement of information that is no longer relevant by newly relevant information. Shifting was interpreted as shifting back and forth between multiple tasks (Allport, Styles, & Hsieh, 1994; Monsell, 1996, 2003). When different tasks (usually choice RT tasks) are mixed within blocks, shifting between tasks typically results in an increase in RT, and a decrease in accuracy (i.e., shift costs). Inhibition was conceptualized as the ability to deliberately inhibit dominant, automatic, or pre-potent responses (Logan & Cowan, 1984).

We expected WCST performance to change during development. That is, we expected set-switching and set-maintenance processes becoming more efficient when children grow older. In addition, we examined the relative contribution of different EF components during development. More specifically, in common with previous studies (e.g., Miyake et al., 2000; Nagahama et al., 2005), we expected that, with advancing age, Shifting plays a significant role in set-switching during WCST performance. Our prediction was based on studies of PFC patients, who show deficient set-maintenance abilities on the WCST. This deficit has been ascribed to the patients' sensitivity to distraction and to the interference of irrelevant information (e.g., Barceló & Knight, 2002; Demakis, 2003; Stuss et al., 2000; see also Konishi, Chikazoe, Jimura, Asari, & Miyashita, 2005).

Previous research revealed that immature inhibitory abilities in young children contribute to deficits in executing efficient strategic behavior (Bjorklund & Harnishfeger, 1990; Harnishfeger, 1995; Miller, 1994). That is, as children mature, inhibitory processes become more efficient and, thus, less irrelevant information enters working memory (see also: Kipp, 2005; Van der Molen, 2000). Here, we hypothesized that Inhibition will contribute most in predicting set-maintenance abilities in the youngest children.

3.2   Method
3.2.1   Participants

The present study included four normal age groups: 51 7-year-olds (30 female, M age = 7.2 (age range = 6-8), 67 11-year-olds (40 female, M age = 11.1 (age range = 10-12), 63 15-year-olds (31 female, M age = 15.4 (age range = 14-16), and 56 young-adults (40 female, M age = 21.0 (age range 18-26). Statistical tests indicated that gender distribution did not differ significantly between age groups, $\chi^2 (3) = 6.09$, $p = .107$.

Children were recruited by contacting regular public local schools; the 21-year-olds were university students, and were recruited through flyers. All children of a particular age group and all students who responded to the flyer were invited to participate in the study. Teachers assisted in the selection process in order to exclude children with learning disabilities, any health problems, neurological damage, or psychiatric problems (as listed in the DSM-IV-TR; APA, 2000). Similar information was derived from a self-report of the 21-year-olds. Informed consent was obtained from parents (for the children) and from adolescents and students. All participants had normal, or corrected-to-normal, vision. The 7- and 11-year-olds received a small present for their participation, the 15-year-olds received € 10, and the young-adults received course credit.

In order to assess their intelligence, participants were administered the Raven Standard Progressive Matrices (SPM; Raven, Court, & Raven, 1985). However, in three 7-year-olds, one 11-year-old, one 15-year-old, and seven 21-year-olds IQ data are missing because these participants were not present during the administration of the SPM. Scores were converted to quartile scores, given the norms of each age group. The Raven quartiles for the 7-, 11-, 15-, and 21-year-olds were 3.7 ($SD = 0.44$), 3.6 ($SD = 0.50$), 3.4; $SD = 0.49$), and 3.7 ($SD = 0.47$), respectively. Statistical tests indicated that there was a significant difference between age groups on Raven SPM percentile, $F (3,221) = 5.29$, $p = .002$. Post-hoc Bonferroni tests indicated significant differences between the young-adults versus the 15-year-olds. Given the significant difference between age groups in Raven SPM quartile scores, we reran the analyses reported below with Raven SPM quartile score as a covariate, and we found no relationship between IQ and the different EF measures (both within-groups and between-groups). The effects of gender and Raven scores on task performance were not further investigated.

3.2.2   Tasks

All participants completed the WCST and the three tasks designed to measure Working Memory, Shifting, and Inhibition. The tasks were taken from a task battery designed to assess EF from childhood through young-adulthood (Huizinga et al., 2006). Task administration was computerized (Toshiba Satellite 1600 laptop; Intel Celeron 800 mHz processor; 15 inch 60 Hz monitor). All tasks required left- and right-hand responses. The response button for the right hand was the "?" key on the computer keyboard, the "z" key served for the left hand response. The WCST was a computerized version of the standard

neuropsychological test. The three tasks used to tap EF component processes were all speeded choice reaction time (RT) tasks. With the exception of the WCST, participants were coached to balance speed and accuracy when responding. Care was taken to ensure that participants understood the instructions, which was verified by verbal report, response accuracy, and stability of the RTs.

Wisconsin Card Sorting Task. We used a computerized version of the WCST (Somsen, Van der Molen, Jennings, & Van Beek, 2000). Against a light-gray background, four key cards, numbered 1 to 4, were presented at the top of the screen. The response cards were taken from the original version of the WCST (Grant & Berg, 1948), and were presented one at a time at the bottom of the screen. The task required participants to match the series of response cards with any of four key cards by pressing the number corresponding to that key card. The display remained visible, until a choice was given. Feedback consisted of a displayed "+" sign, following a correct response, and a "-"sign, following an incorrect response.

Response cards could be matched on color (red, green, blue, yellow), shape (triangle, star, cross, circle), or number (1, 2, 3, 4). Once the participant made 10 consecutive correct sorts, the sorting principle was altered. The task was terminated either after the participant completed 6 categories (e.g., shape, color, form, color, form, shape), or after the maximum of 128 trials was reached. The order of the sorting principles was randomized, with the constraint that the same sorting principles did not occur consecutively. The test was administered according to the procedure outlined in the Heaton manual (Heaton et al., 1993). Three following variables of interest were selected. Perseverative error responses were defined as errors resulting from persistence in responding to a stimulus characteristic that is no longer correct (Barceló & Knight, 2002; Heaton et al., 1993). Efficient error responses refer to a switch to the wrong category on the second trial on an otherwise clear series (i.e., series with no further errors other than the first warning error). Efficient errors were scored only in the 2nd trial of the series and were incompatible with any other error in the remaining trials of that series (Barceló & Knight, 2002). Distraction errors were defined as a switch to the wrong category different from the one chosen in the previous trial (Barceló & Knight, 2002). The proportions of the respective error types were calculated by computing the number of a particular error type relative to the number of trials administered, multiplied by 100. In addition, we examined the number of correct responses (i.e., the number of responses that match the correct sorting principle in effect at the time the response is made) and the number of categories achieved (i.e., the number of sequences of 10 consecutive correct matches to the criterion sorting strategy). The proportion of correct responses was indexed by computing the number of correct responses relative to the number of trials needed to complete the test, multiplied by 100.

Working Memory. In the Tic Tac Toe task (adapted from Milner, 1971), participants were required to retain visual information about the orientation of a pattern of figures active in their working memory. The task consisted of a memorizing phase and a recognition phase. A pattern consisting of *X*s and *O*s was presented within a 3x3 grid during the memorization phase. Working memory load was varied by using patterns consisting of three vs. four letters (i.e., low memory load vs. high memory load). The recognition phase was initiated by pressing the space bar. During this phase, the *X*s and *O*s were presented one after another at different positions in the grid, each for a period of 600 ms, in series that varied from four to seven presentations for the low memory load, and from four to nine presentations for the high memory load. As soon as the pattern of *X*s and *O*s included the pre-specified pattern indicated in the memorization phase, participants were required to press a button. Participants first received a practice block of 3 trials, followed by one block including 15 trials for low memory load and one block including 15 trials for high memory load. The low and high memory load blocks were counterbalanced across participants to control for order effects. Participants had 3500 ms to respond; the time interval between trials varied randomly between 900 and 1100 ms (drawn from a uniform distribution). The main dependent variables were the accuracy proportions and median RTs on the low and high memory load blocks.

Shifting. In the Local-Global task (adapted from Miyake et al., 2000) participants responded to randomly presented rectangles or squares by pressing a left or right response button, respectively. Larger (global) rectangles/squares consist of smaller (local) rectangles or squares. In one block participants responded to the local figure, in a second block they responded to the global figure (blocks 1 and 2, in randomized order; 30 practice trials and 50 experimental trials per block), and in the third block they alternated between a series of four "local" trials and a series of four "global" trials (block 3; 90 practice trials, 150 experimental trials). A cue indicated to which dimension (global or local) the participants should respond. Cues relating the global (local) dimension consisted of a big (small) square, which was presented on one side of the target stimulus, and a big (small) rectangle, which was presented on the other side of the target stimulus. The color of cues and target was red. They remained on the screen until a response was given. Participants had 3500 ms to respond. The time interval between presentation of the cue and of the target stimulus was 500 ms. The interval between the response and the presentation of the cue was fixed at 1000 ms. The main dependent variables were the accuracy proportions, and median RTs on task repetition and task alternation trials.

Inhibition. In the present version of the Stop-signal task (adapted from Van Boxtel, Van der Molen, Jennings, & Brunia, 2001), participants had to respond as fast as possible to a left or right pointing arrow by a left or right button press. On 25% of the trials, the color of the arrow changed unpredictably from green to red, indicating that the response to the arrow stimulus should be inhibited. The time interval between arrow

onset and arrow color varied depending on the participant's performance. A dynamic tracking algorithm was used to ensure that stopping approximated 50% correct inhibited responses. The stimulus remained on the screen until a response was given. Participants had 1250 ms to respond. The time interval between the response and the arrow onset on the subsequent trial varied randomly between 1650 and 2150 ms (drawn from a uniform distribution). There were 50 practice trials and two blocks of 100 experimental trials. The main dependent variable were the proportion of correctly inhibited trials and the median stop stimulus reaction time (SSRT), reflecting the latency of the internal response to the stop signal (see Logan, 1994).

3.2.3   Procedure

The order of tasks was counterbalanced across participants. The WCST, however, was always administered last, in view of the inter-individual variation in total time needed to complete the task. Both at the schools and the university, the tasks were administered in a dimly-lit quiet room. There were 3 minute-breaks between tasks, and a 10-minute break after 3 tasks. Each test session lasted approximately 1.5 hours. At the end of the test-session, the 15-year-olds and 21-year-olds completed a paper-and-pencil version of the Raven SPM (Raven et al., 1985); the children completed this task individually in the classroom (with all participants present).

3.2.4   Exclusion criteria

A participant was excluded if one of the following criteria was met: Failure to complete at least one category at the WCST; mean accuracy percentage at the low memory load of less than 55% (Tic Tac Toe task); less then 55% correct on the first two blocks (Local-Global task); proportion correct inhibits lower than 20% or higher than 80% (Stop Signal task). Sixteen 7-year-olds, six 11-year-olds, and two 21-year-olds were excluded (this exclusion did not result in a significant change in gender distribution and Raven score per age group)[1].

3.3   Results

The goal of the present study was to examine the development of WCST performance by integrating the approaches advanced by Barceló and Knight (2002) and by Miyake et al. (2000). We scored WCST performance in terms of set-switching and set-maintenance abilities (see Barceló & Knight (2002). Subsequently, we examined the relative

---

[1] The relatively large amount of 7-year-olds excluded from the study was largely caused by failure on the Tic Tac Toe test (9 participants).

contributions to these abilities of three EF components Working Memory, Shifting, and Inhibition, as distinguished by Miyake et al. (2000). The results are presented in three sections. The first section focuses on the performance on the WCST and the three EF components tasks. This allows for the comparison of results obtained in the present sample to those reported in the literature. The second section focuses on the correlations between the measures reflecting set-switching and set-maintenance processes on the WCST. The third section focuses on the extent, to which EF components Working Memory, Shifting, and Inhibition tasks predict WCST set-switching and set-maintenance. In the analyses reported below, Age Group (7-year-olds, 11-year-olds, 15-year-olds, 21-year-olds) was included as between-subjects factor.

3.3.1   Developmental trends on the WCST and EF component tasks

WCST. The means and standard deviations of the four age groups are reported in Table 1 (see also Figure 1; left panel: errors related to set-switching; right panel: errors related to set-maintenance). A MANOVA performed on the proportions of correct responses, the number of categories achieved, and the proportions of perseveration and distraction errors revealed significant main effects for Age Group (p's $<.001$, $\eta_p^2 = .16$, $\eta_p^2 = .14$, $\eta_p^2 = .10$ and $\eta_p^2 = .11$, respectively). A trend was observed for the main effect of Age Group on the proportion of efficient errors, F (3,217) = 2.60, p = .052, $\eta_p^2 = .04$. Post-hoc Bonferroni analyses showed that the proportion of correct responses was smaller in 7-year olds than in 11-year olds (ns.)[2], and smaller in 11-year olds than in 15-year olds. The 15-year olds did not differ from 21-year olds. The number of categories completed was smaller in 7-year olds than in 11-year olds, smaller in 11-year olds than in 15-year olds, and smaller in 15-year olds than in 21-year olds. In addition, the proportion of perseveration errors was larger in 7-year-olds than in 11-year-olds. The 11-year-olds did not differ from 15-year-olds, who did not differ from 21-year-olds. The age groups did not differ from each other with respect to efficient errors. The proportion of distraction errors was largest in 7-year-olds than in 11-year-olds, larger in 11-year-olds than in 15-year-olds, and larger in 15-year-olds than in 21-year-olds.

Thus, we observed two distinct developmental trajectories for set-switching and set-maintenance processes. That is, set-switching performance (i.e., perseveration errors) improved until adolescence, whereas set-maintenance performance (i.e., distraction errors) improved until young-adulthood.

---

[2] One anonymous reviewer argued that the non-significant difference between 7- and 11-year-olds resulted from the fact that a large number of 7-year-olds was excluded due to their failure on the Tic Tac Toe task. As a result of this exclusion 7-year-olds with relatively better working memory abilities might be included in the study. Subsequent Bonferroni analysis that included the 7-year-olds that were initially excluded from the Tic Tac Toe task did however not reveal a significant difference on "categories achieved" between 7- and 11-year-olds (p = .43).

	7-year-olds	11-year-olds	15-year-olds	21-year-olds
Correct responses (%)	47.9 (17.3)	57.7 (17.6)	66.1 (14.5)	65.6 (15.1)
Categories achieved (#)	3.1 (1.7)	3.7 (1.9)	4.7 (1.7)	4.8 (1.8)
Perseverative errors (%)	24.8 (13.0)	18.3 (11.5)	15.8 (9.7)	14.8 (9.6)
Efficient errors (%)	.4 (.5)	.5 (.8)	.7 (.8)	.8 (1.1)
Distraction errors (%)	22.3 (10.9)	20.5 (9.5)	14.7 (5.1)	16.5 (7.6)

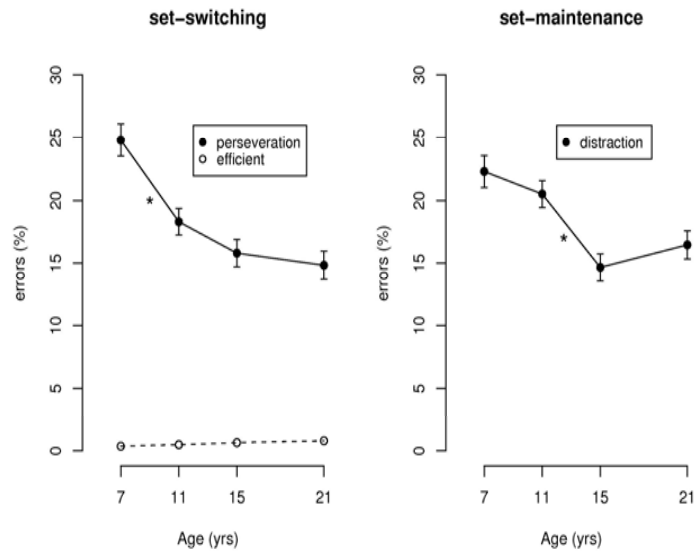*Table 1. WCST dependent variables per Age Group; SDs between parentheses*

*Figure 1. Error rates (%) associated with WCST performance; * significant difference relative to the subsequent age group*

Working Memory. The first response in each block and trials with RTs shorter than 200 ms were excluded from the analysis. Separate repeated measures ANOVAs were performed, with Load (low vs. high) as a within-subjects factor. Dependent variables were response speed (median RTs on correct trials) and the square root of error

proportions. ANOVAs, performed on accuracy (% correct) and RT, revealed significant main effects of Age Group and Load (p's < .001). In addition, for both accuracy and RT, the interaction of Age Group and Load was significant, $F(3,217) = 29.06$, $p < .001$, $\eta_p^2 = .29$ and $F(3,217) = 3.85$, $p = .010$, $\eta_p^2 = .05$, respectively (see Figure 2).
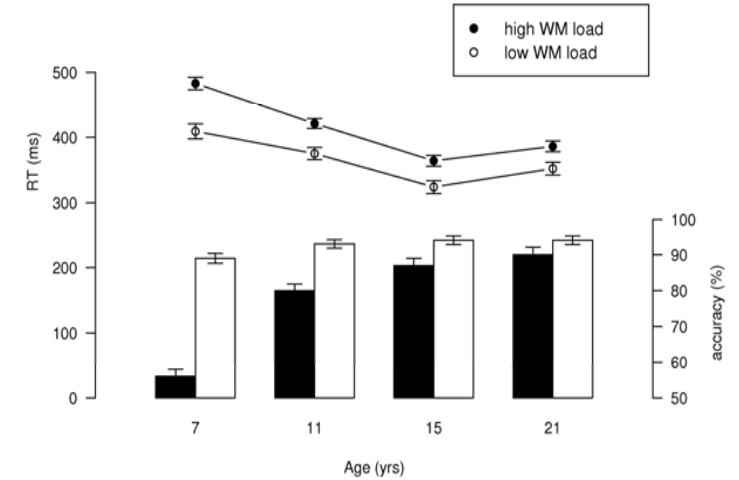
Figure 2. Response latencies (ms) and accuracy (%) associated with Tic Tac Toe task performance, per Age Group

A MANOVA performed on the difference scores (i.e., high load vs. low load) indicated that in all age groups accuracy of performance decreased with the increase in working memory load. This effect was most pronounced in younger children, and decreased with age. Post-hoc Bonferroni analyses revealed that the effect on accuracy differed significantly between the 7- and 11-year-olds, 11-year-olds did not differ from 15-year-olds, who did not differ from 21-year-olds. In addition, the MANOVA did not show RT differences between 7-year-olds and 11-year-olds, 11-year-olds and 15-year-olds, and 15-year-olds and 21-year-olds.

Shifting. The first response in each block, trials with RTs shorter than 200 ms and trials that were preceded by error trials were excluded from the analysis. Separate repeated measures ANOVAs were carried out with Trial Type (repetition vs. alternation) as within-subjects factor. ANOVAs performed on the square root of accuracy (% correct)

and RT revealed a significant main effect of Age Group (p's < .001). The main effect of Trial Type failed to reach significance on accuracy (p = .09), but the main effect of Trial Type on RT was significant (p < .001). Importantly, for both accuracy and RT, the interaction of Age Group and Trial Type was significant, F (3,217) = 3.47, p = .017, η_p^2 = .05 and F (3,217) = 16.27, p < .001, η_p^2 = .18, respectively (see Figure 3).
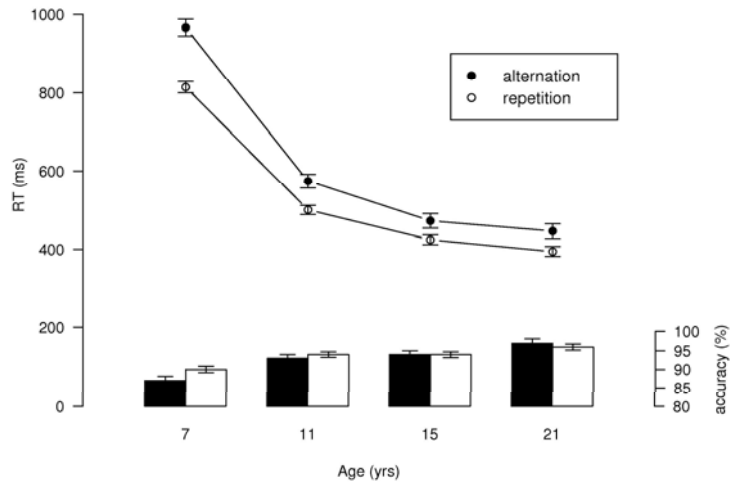

Figure 3. Response latencies (ms) and accuracy (%) associated with the performance on the mixed block of the Local-Global Shifting task performance, per Age Group

A MANOVA performed on the shift costs difference scores (i.e., the increase in RT or accuracy in alternation trials compared to repetition trials within mixed-task blocks). The MANOVA did not show accuracy differences between 7-year-olds and 11-year-olds, 11-year-olds and 15-year-olds, and 15-year-olds and 21-year-olds. In addition, we found in all age groups shift costs in terms of RT were evident. This effect was most pronounced in children, and decreased with age. Shift costs in terms of RT differed significantly between the 7- and 11-year-olds, whereas 11-year-olds did not differ from 15-year-olds, who did not differ from 21-year-olds.

Inhibition. The first four responses in each block were excluded from the analysis. A MANOVA was performed on response ratio (%) and SSRT. Participants were able to

stop their responses on stop-signal trials in about half of the trials on which a stop signal was present (M = 48.9%, SD = 8.2), and the MANOVA revealed no significant effect of Age Group, F (3, 217) = 1.38, p = .251, η_p^2 = .02. This finding indicates that the tracking algorithm worked well in all age groups. For SSRT, a significant main effect of Age Group was observed, F (3,217) = 18.58, p < .001, η_p^2 = .20, showing that SSRT decreased from 297 ms (SD = 89) ms in 7-year-olds, through 232 (SD = 57) ms in 11-year-olds and 214 (SD = 67) in 15-year-olds, to 203 (SD = 55) ms in 21-year-olds. Post-hoc Bonferroni analysis showed that the speed of response inhibition increased during childhood, as indicated by a significant difference between the 7-year-olds and the 11-year-olds, whereas 11-year-olds did not differ from 15-year-olds, who did not differ from 21-year-olds.

3.3.2 Correlations between set-switching and set-maintenance on the WCST

To explore the relation between set-switching and set-maintenance processes during WCST performance, principal components analyses (PCA) with orthogonal rotation were performed per age group (see Tables 2 and 3). To determine factor consistency, a loading of .40 was used as a criterion. Using the criterion of Eigenvalues greater than one, two factors were extracted in the youngest children, and one factor in the three oldest age groups.

Age Group	Component	Eigenvalue	% of variance	Cumulative %
7-year-olds	1	1.467	49.90	48.90
	2	1.117	37.23	86.13
11-year-olds	1	1.745	58.18	58.18
15-year-olds	1	1.484	49.47	49.47
21-year-olds	1	1.816	60.55	60.55

Table 2. Eigenvalues and variance (%) accounted for by each Component per Age Group

In the 7-year-olds, nature of the two factors was as follows: Factor 1 included the proportion of perseverative errors and the proportion of efficient errors (i.e., representing a set-switching factor), whereas Factor 2 included the proportion of distraction errors (i.e., representing a set-maintenance factor). This distinction could not be made in the three oldest age groups resulting in a single factor extracted by the PCA.

Variable	Age Group	Factor	
		1	2
	7-year-olds		
Perseverative errors		.837	
Efficient errors		-.871	
Distraction errors			.960
	11-year-olds		
Perseverative errors		.747	
Efficient errors		-.838	
Distraction errors		.698	
	15-year-olds		
Perseverative errors		.763	
Efficient errors		-.477	
Distraction errors		.822	
	21-year-olds		
Perseverative errors		.776	
Efficient errors		-.786	
Distraction errors		.772	

Table 3. Factor loadings for the proportions of perseverative errors, efficient errors, and distraction errors per Age Group

3.3.3 Prediction of WCST performance

We examined the relative contribution of Working Memory, Shifting, and Inhibition to the factor structure that emerged from the PCA. The mean accuracy score on the Tic Tac Toe task, the median RT on alternation trials on the Local-Global task, and SSRT were the indicators of Working Memory, Shifting, and Inhibition, respectively. The factor scores obtained in the PCA reported above were the criterion variables in backward regression analyses with a probability of F-to-remove > .10.

The regression analyses revealed that in 7-year-olds Factor 1 was best predicted by the model including Shifting, $F (1, 42) = 3.05$, $p = .088$, explaining 6.9% of the variance. Factor 2 was best predicted by the model including Inhibition, $F (1, 42) = 3.58$, $p = .066$, explaining 8.0% of the variance. In 11-year-olds Factor 1 was best predicted by the model including Shifting, $F (1, 62) = 7.66$, $p = .007$, explaining 11.2% of the variance. In addition, in 15-year-olds Factor 1 was best predicted by the model including Shifting and Working Memory, $F (2, 58) = 4.63$, $p = .014$, explaining 14.2% of the

variance. Finally, in 21-year-olds Factor 1 was best predicted by the model including Working Memory, $F (1, 55) = 7.58$, $p = .008$, explaining 12.3% of the variance.

3.4 Discussion

The present study was designed to investigate the development of set-switching and set-maintenance processes underlying WCST performance (see also Barceló & Knight, 2002). In addition, we examined the relative contribution of the EF components Working Memory, Shifting, and Inhibition to set-switching and set-maintenance during WCST performance (see also Miyake et al., 2000).

3.4.1 Developmental trends of set-switching and set-maintenance on the WCST

Barceló and Knight (2002) distinguished between set-switching and set-maintenance processes involved in WCST performance. Set-switching abilities are associated with efficiently switching to the new sorting rule on the basis of feedback, and are indexed by the proportions of perseverative and efficient errors. In this study, performance of set-switching abilities (i.e., perseverative errors) reached young-adult levels in 11-year-olds. This finding is consistent with the results of prior developmental studies (Chelune & Baer, 1986; Crone et al., 2004; Lehto, 2004; Welsh et al., 1991). It should be noted, however, that in contrast to other studies (Barceló & Knight, 2002; Crone et al., 2004), the number of efficient errors observed here was rather low. Efficient errors occur when a participant switches to the wrong sorting rule in the second trial of an otherwise clear series comprising the new sorting rule (i.e., series with no errors other than the first error indicating that the sorting principle changed; Barceló & Knight, 2002). We hypothesize that the current finding of relatively few efficient errors is due to the absence of ambiguous trials in other studies. The current computerized implementation of the standard WCST included ambiguous trials, which reduced the chances of committing efficient errors.

Set-maintenance requires the retention in mind of the current sorting rule throughout varying stimulus conditions, while ignoring irrelevant aspects of the stimuli, and is indexed by the proportion of distraction errors. In the present study, set-maintenance performance on the WCST continued to develop into adolescence. This finding is consistent with the results of other studies, reporting the improvement of set-maintenance performance into adolescence (e.g., Chelune & Baer, 1986; Crone et al., 2004). This developmental trend indicates that with advancing age children are less susceptible to random failures to maintain set during WSCT performance. The reported age-related increase in set-maintaining performance might reflect an increased ability to

keep information on-line (Barceló & Knight, 2002; Case, 1992; Crone et al., 2004; Gathercole et al., 2004).

3.4.2   Developmental trends on EF component tasks

In the present study, we observed performance relating to Working Memory to develop into adolescence, whereas Shifting and Inhibition performance reached young-adult levels in 11-year-olds. The developmental trend for Working Memory is similar to the age-group differences found in earlier studies using the Corsi blocks test (a test with a similar format as the Tic Tac Toe test). These studies showed that adult levels of performance are not reached until adolescence (Hitch, Halliday, Dodd, & Littler, 1989; Issacs & Vargha-Khadem, 1989; Kemps, De Rammelaere, & Desmet, 2000; Wilson, Scott, & Power, 1987).

Similar to other studies, Shifting abilities (as reflected by shift costs, i.e., the difference in RT and accuracy on alternation trials relative to repetition trials) reached young-adult levels of performance by the age of 12. The larger shift costs in younger children have been interpreted to reflect immature levels of executive function (Cepeda et al., 2001; Zelazo, Craik, & Booth, 2004), or a delay in the retrieval of S-R links from memory (Crone, Bunge et al., 2006).

Finally, consistent with previous studies, performance on the stop-signal inhibition task reached young-adult levels of performance by the age of 12 (Van den Wildenberg & Van der Molen, 2004; Williams et al., 1999). This finding is consistent with the literature assuming that the ability to inhibit develops rapidly during childhood (e.g., Dempster, 1992; for reviews see also Kipp, 2005; Van der Molen, 2000).

Thus, performance analyses showed different developmental trends in the performance on Working Memory, Shifting, and Inhibition tasks. The performance on the Working Memory task continued to develop into young-adulthood whereas the performance on the Shifting and Inhibition tasks attained mature levels during adolescence. The finding of different developmental trends is consistent with previous studies reporting that EF becomes more efficient as children grow older, and that adult-levels of performance on various EF tasks are attained at different ages during childhood and adolescence (e.g., Cepeda et al., 2001; Huizinga & Van der Molen, 2006; Luna et al., 2004; Van den Wildenberg & Van der Molen, 2004). Importantly, these findings are usually interpreted in terms of PFC maturation (e.g., Gogtay et al., 2004; Sowell et al., 2004). That is, the improvement of EF during childhood and adolescence has been interpreted to reflect gradually less diffuse and more focal PFC activation with advancing age (e.g., Amso & Casey, 2006; Casey et al., 2005).

3.4.3   Contribution of EF component processes to WCST performance

PCA revealed that, in 7-year-olds, set-switching and set-maintenance abilities loaded on two respective factors. This pattern of findings was taken to suggest that successful WCST performance of 7-year-olds draws upon both the genuine ability to flexibly switch to the new sorting principle and the ability to maintain set. Interestingly, in the older age groups, a separate set-maintenance factor was not observed, suggesting the development of the ability to resist distraction from irrelevant interfering information. This finding is consistent with previous literature, in which the ability to maintain set is associated with an increased ability to keep information on-line in working memory (Barceló & Knight, 2002; Case, 1992; Crone et al., 2004). The finding of WCST performance being represented by two factors in 7-year-olds and by only one single factor from late childhood on is taken to suggest the increasing efficiency of EF, most likely reflecting the increasing focal PFC activation that is observed during childhood and adolescence (e.g., Amso & Casey, 2006; Casey et al., 2005).

We regressed WCST performance (as indexed by the respective principal component scores reflecting the Barceló and Knight (2002) scoring method) on the Working Memory, Shifting, and Inhibition tasks. The regression analyses revealed that the set-switching factor observed in 7-year-olds was predicted by Shifting (as expected) and the set-maintenance factor by Inhibition (as expected). The finding that Shifting contributed to WCST performance is in accordance with previous studies showing that the ability to flexibly shift between tasks contributes to WCST performance (e.g., Miyake et al., 2000; Nagahama et al., 2005). The finding that Inhibition is a good predictor of the set-maintenance factor is in accordance with theories assuming that young children are more susceptible to interference from irrelevant information (e.g., Dempster, 1992; Ridderinkhof & Van der Molen, 1995; for reviews see also: Kipp, 2005; Van der Molen, 2000). Distraction errors can be regarded as an untimely reset of the contents of working memory due to an inadequate ability to refrain from processing interfering stimuli or to inhibit interfering responses (Barceló, 1999). The ability to deal with interference fosters the active selection of relevant task information. Specifically, as shown in the current study, young children are less likely to inhibit responses on the WCST that were previously correct, but currently incorrect. That is, response selection on the current trials is guided by a rule that was correct on previous trials but incorrect on the current trial.

The single factor found in the three oldest age groups was best predicted by Shifting (in the 11-year-olds), Shifting and Working Memory (in the 15-year-olds), and by Working Memory (in the 21-year-olds). Thus, in 11- and 15-year-olds successful WCST performance appears to draw upon the ability to flexibly switch to the new sorting rule. The finding of Working Memory involvement in WCST performance of 11- and 15-

year-olds replicates earlier work with adults (e.g., Bechara et al., 1998; Huizinga et al., 2006; Konishi et al., 1999). Finally, it was found that from late childhood on, Inhibition does not predict WCST performance. This finding is taken to suggest to reflect the complete maturation of inhibitory processes by around age 12, as reported earlier in the literature (Bédard et al., 2002; Bunge et al., 2002; Durston et al., 2002; Ridderinkhof & Van der Molen, 1995; Van den Wildenberg & Van der Molen, 2004). The current results suggest that, with advancing age, one is increasingly able to keep different alternatives (rules) in mind, while searching for the new correct sorting rule on the basis of feedback. This notion is supported by studies reporting the emergent ability to update information in working memory with advancing age (e.g., Beveridge et al., 2002; Brocki & Bohlin, 2004; DeLuca et al., 2003; Gathercole et al., 2004; Luciana et al., 2005; Luna et al., 2004; see also Barceló & Knight, 2002; Case, 1992; Crone et al., 2004). Interestingly, this notion is supported by the developmental trend into adolescence of Working Memory observed in the current study.

3.4.4   Conclusion

The current study set out to assess the developmental trajectories of set-switching and set-maintenance on the WCST. In addition, the relative contribution of the EF components Working Memory, Shifting and Inhibition was examined to the observed age-group differences in WCST performance. The major findings that emerged from this study were (a) young-adult levels of performance were reached in 11-year-olds for set-switching, and in 15-year-olds for set-maintenance, (b) set-switching and set-maintenance loaded on two factors in the 7-year-olds but on a single factor in the other age groups, and (c) regression analyses yielded differential contributions of Working Memory, Shifting and Inhibition to set-switching and set-maintenance on the WCST during development.

The current pattern of findings is interpreted to suggest distinct developmental trends in set-switching and maintenance abilities required by the WCST. It could be argued here that these interpretations must be qualified in view of the relatively small sample of participants. It should be noted, however, that the amount of variance explained in all age groups is hardly cause for alarm (i.e., it ranged from about 55% in the 11- and 15-year-olds to 87% in the 7-year-olds). Moreover, the current findings are in accord with previous work showing developmental improvement in set-switching and set-maintenance in WCST performance (Crone et al., 2004; see also Chelune & Baer, 1986; Chelune & Thompson, 1987; Huizinga et al., 2006; Levin et al., 1991; Welsh et al., 1991). In addition, previous neuro-imaging studies showing distinct developmental trajectories of EF components support the findings reported in this study (Bunge et al., 2002; Crone, Donohue, Wendelken, Honomichl, & Bunge, 2006; Klingberg, Forssberg, & Westerberg, 2002; Luna & Sweeney, 2004). Within a broader context, the different

developmental patterns found in this study reflect differential development of PFC contributions to EF (for reviews see Casey et al., 2005 and Diamond, 2002). Future research is needed to provide further insight in the developmental pattern of PFC activation in relation to the development of processes underlying WCST performance.

Chapter 4 ● Development of switching from color to shape, and from stopping to going

Abstract

This study examined developmental differences in the ability to switch between tasks and to shift between stopping and going. Three age groups (6-7 year-olds, 11-12 year-olds, and 20-21 year olds) performed on a hybrid task, in which participants were required to consider the color or shape of a target stimulus, and to execute a choice response on some trials, but a disjunctive response on others. The paradigm allowed the assessment of the speed of choice repetitions and choice alternations (i.e., switches from color- to shape-responses or vice versa, and shifts from go- to choice-responses or from nogo- to choice-responses) and of disjunctive repetitions (i.e., from go- to go-responses) and disjunctive alternations (i.e., shifts from nogo- to go-responses and from choice- to go-responses). The results showed that the costs involved in switching between choice responses decreased with advancing age into adolescence. Similarly, the costs in shifting from stopping to going decreased with the children's age, but already reached mature levels by late childhood. The results indicate that switching between color and shape responses and shifting between stopping and going are both time-consuming processes that mature at different rates.

4.1 Introduction

A rapidly growing literature has demonstrated that executive function (EF) improves as children grow older (Beveridge, Jarrold, & Pettit, 2002; Casey, Giedd, & Thomas, 2000; Diamond, 2002; Diamond, Briand, Fossella, & Gehlbach, 2004; Klenberg, Korkman, & Lahti Nuuttila, 2001; Lehto, Juujaervi, Kooistra, & Pulkkinen, 2003; Luciana & Nelson, 1998; Welsh, 2002; Welsh, Pennington, & Groisser, 1991). In a broad sense, EF refer to general-purpose cognitive control mechanisms that regulate behavior to achieve a future goal. More specifically, EF refers to the many skills that are required to prepare and to execute complex behaviors, including planning, inhibition, flexibility, organization, self-monitoring, and mental representation (e.g., Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Duncan & Owen, 2000; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Stuss, Shallice, Alexander, & Picton, 1995). EF is considered effective if it facilitates the individual's flexible adjustment to the changes in the environment. Young children are often characterized as stimulus bound and impulsive (cf. Zelazo, Craik, & Booth, 2004), and thus are likely to perform poorly on tasks that require flexible switching between task demands, and the inhibition of stimuli and responses that are no longer relevant (Cepeda, Kramer, & Gonzalez de Sather, 2001; Chelune & Baer, 1986;

Ridderinkhof, Van der Molen, Band, & Bashore, 1997; Welsh et al., 1991; Zelazo, Muller, Frye, & Marcovitch, 2003).

The Wisconsin Card Sorting Task (WCST) has provided the main paradigm in the study of development of the ability to adjust to changing task demands (Berg & Grant, 1948; Heaton, Chelune, Talley, Kay, & Curtis, 1993; Milner, 1963). The WCST requires individuals to sort cards on the basis of the number, shape, and color of geometric objects that are printed on the cards. The sorting rule has to be derived from feedback provided by the tester. Following 10 correct sorts, the sorting rule is changed without warning, and the individual must again derive the new sorting rule on the basis of feedback. A number of developmental studies, either using the original task or a computerized version of it, demonstrated that, with advancing age, children need fewer trials to discover the first sorting rule, make less perseverative responses when the sorting rule is changed, and thus solve more categories (Chelune & Baer, 1986; Huizinga, Visser, & Van der Molen, 2004; Welsh et al., 1991). These findings are usually interpreted in terms of developmental improvement in EF, which in turn is associated with the maturation of the prefrontal cortex (e.g., Casey et al., 2000; Chugani, Phelps, & Mazziotta, 1987; Giedd et al., 1999; Huttenlocher, 1979; Pfefferbaum et al., 1994; Yakovlev & Lecours, 1967; for reviews, see Diamond, 2002; Van der Molen & Ridderinkhof, 1998).

Research based on the WCST supports the hypothesis that EF during childhood improves with age. However, the WCST is a complex task that requires a variety of component control processes, including problem solving and performance monitoring, to discover the new rule after a change. In addition, WCST performance requires working memory and the ability to flexibly switch responses. In view of this complexity, several authors advocated the use of tasks, designed specifically to assess cognitive flexibility and the ability to inhibit inappropriate responses (Cepeda et al., 2001; Stuss et al., 1995). They identified the "task-switching" paradigm as a valuable tool for assessing component processes, which underlie cognitive flexibility in the absence of problem solving (cf. Cepeda et al., 2001). The task-switching paradigm requires the individual to switch between two simple tasks, such as deciding whether the color of the stimulus is red or blue, or deciding whether the shape of the stimulus is a circle or a triangle. The two tasks are presented in mixed blocks, which allows for a comparison of performance on task repetitions and task alternations. Typical responses are observed to be slower and less accurate on alternation trials, than they are on repetition trials (e.g., Allport, Styles, & Hsieh, 1994; De Jong, 2000; Jersild, 1927; Meiran, 1996; Rogers & Monsell, 1995; Spector & Biederman, 1976). The differences in performance on alternation and repetition trials are coined "switch costs". These are attributed to the required reconfiguration of task sets in alternation trials (De Jong, 2000; Meiran, 1996; Rogers & Monsell, 1995), to the passive decay of the previous task set from working memory (e.g., Meiran, Chorev, & Sapir, 2000; Spector & Biederman, 1976), and to proactive

interference (Allport et al., 1994; Wylie & Allport, 2000). The task-switching paradigm has been used extensively in research into EF in adults. This research has established that switching costs are affected by many factors. Consequently, there is much speculation regarding the nature of the processes, that are involved in adult task-switching (see Monsell, Yeung, & Azuma, 2000; Rubinstein, Meyer, & Evans, 2001 for reviews of task-switching literature).

Although the task-switching paradigm has proven to be a productive vehicle to assess cognitive flexibility in adulthood, to day only a few studies employed this paradigm to assess developmental change. All these studies, bar one (Crone, Bunge, Van der Molen, & Ridderinkhof, 2006), adopted a lifespan orientation, which involves the comparison of multiple age groups. These studies however did not statistically compare children and adults, or compared only one child group relative to young adults. Cepeda et al. (2001), for example, used a task-switching paradigm to assess lifespan change in cognitive flexibility from age 7 to 82 years. They presented their participants with either one or three 1's or 3's (i.e., 1, 111, 3, or 333), and required them to determine whether the number 1 or 3 was presented on the screen (task A), or whether a single number or three numbers were displayed on the screen (task B). The two tasks were presented in mixed blocks, and a cue was given to inform the participants about the task required on a particular trial. The results of their study suggested that (i) switch costs decreased from childhood to young adulthood, and (ii) switch costs decreased when the time-interval between cue and stimulus was lengthened, with children showing larger benefits of preparation time than adults.

Span, Ridderinkhof and Van der Molen (2004) conducted a lifespan study, in which they presented schematic faces to the participants, who were required to decide whether the face was female or male, whether the face was happy or sad, or whether spectacles were present of absent. The three tasks were presented in mixed blocks such that tasks repetitions and tasks alternations occurred with equal frequencies. In line with the findings reported by Cepeda et al. (2001), Span et al. observed that switching costs were considerably larger in children (age range 6-10 years) than in young adults, even when controlling for group differences in basic processing speed (i.e., the speed of responding on a simple reaction task). This finding was interpreted in support of a developmental increase in cognitive flexibility.

Another lifespan study, reported by Zelazo et al. (2004), obtained a similar result. In this study, participants were required to determine the color of several stimuli (color task), or stimulus shape (shape task). The tasks were cued with one cue occurring on 80% of the trials, and the other cue on the remaining 20% of the trials. In one condition, the task cue was presented visually, and in the other condition it was presented auditory. In both conditions, 8-year-olds were found to commit more perseverative errors than young adults. This finding was taken to suggest that cognitive flexibility increases during

childhood. Contrasting results were obtained in a recent lifespan study reported by Kray, Eber, & Lindenberger (2004). They asked participants to determine the category to which a stimulus belonged (animal or fruit), and its color (gray or not-gray). The object and color tasks were presented in mixed blocks with a fixed interval between the task cue and the stimulus. In contrast to the previous studies, Kray et al. observed that the costs involved in switching tasks did not discriminate between 9 year-olds and young adults.

Finally, Crone and colleagues (2006) conducted a developmental study of task-switching, in which 8-year olds, 11-year-olds, and young adults were required to respond to a stimulus, which appeared to the left or to the right on a screen, by pressing either a spatially compatible or a spatially incompatible response button, depending on the shape and color of the stimulus. Crone et al. focused on the switch costs associated with response repetitions and alternations, in the spatially compatible and incompatible tasks. They observed that switch costs decreased with increasing age of the children. Consistent with the adult literature, Crone and colleagues obtained a highly significant interaction between the effects of task-switching and response repetition/alternation (Kray & Lindenberger, 2000; Meiran, 1996; Rogers & Monsell, 1995). That is, a benefit was enjoyed when the response was repeated on non-switch trials (see also Kirby, 1980; Luce, 1986; Soetens, Boer, & Hueting, 1985), but a cost was incurred when the response was repeated on switch trials. Interestingly, these costs were much larger in children than adults. Crone et al. interpreted these results as indicating that children's switch costs are, at least in part, due to a slow decay in residual memory traces associated with the previous trial. This slow decay was thought to interfere with children's ability to switch to currently intended actions (cf. Crone et al., 2006). This interpretation received support from the additional finding showing that the age-related trend in switch costs decreased with the increase of the response-to-stimulus interval from 50 to 500 ms.

4.1.1   The present study

The present study was conducted to further explore developmental differences in task-switching. By extending the task-switching paradigm, we will assess switching between tasks and switching between response execution, and response inhibition. It is well known that switching from responding to non-responding is difficult (e.g., Woodworth & Schlosberg, 1954). Thus, it is difficult to inhibit the prepared response when the stimulus happens to be a nogo signal (for a review, Luce, 1986), or when the go signal is overruled by a sudden stop-signal (for a review, Logan, 1994). Developmental studies have indicated repeatedly that, compared to young adults, children experience more difficulty in withholding their response to nogo signals (e.g., Jonkman, Lansbergen, & Stauder, 2003) and stop-signals (e.g., Bédard et al., 2002; Van den Wildenberg & Van der Molen, 2004; Williams, Ponesse, Schachar, Logan, & Tannock, 1999). Conversely, it is difficult

to shift from non-responding to responding. Several studies have indicated that responses on a non-signal trial are considerably slower when the trial is preceded by a stop-signal trial rather than another non-signal trial (e.g., Kleinsorge & Gajewski, 2004; Rieger & Gauggel, 1999; Rieger, Gauggel, & Burmeister, 2003; Schuch & Koch, 2003). Likewise, the speed of responding is slower when shifting from a nogo trial to a choice reaction trial, in comparison to the speed of responding on a choice trial that is preceded by another choice reaction trial (Hoffmann, Kiesel, & Sebald, 2003). Employing a stop-change paradigm, Band and co-workers observed that, compared to adults, young children needed disproportionally more time to switch from the prepared response to an alternate response, when a change signal, which accompanied the respond signal, required them to do so (Band, Van der Molen, Overtoom, & Verbaten, 2000). The aim of the present study is to determine whether development of switching between tasks, and shifting between responding and inhibition, proceed along a similar course.

The task-switching paradigm used in the present study is a hybrid version of the "alternating-runs" paradigm and the "task-cue" paradigm. This version was used so as to reduce the demands on working memory as much as possible. Two tasks were used in which the requirement was to respond to stimulus color or to respond to stimulus shape.

In the alternating-runs paradigm, the task changes predictably every nth trial, and the participant is assisted in keeping track of the current task by a cue, such as the position of the stimulus on the screen (e.g., Rogers & Monsell, 1995). In the present paradigm (see Figure 1), the requirement to consider color (or shape) changed every 5th trial, and the requirement to switch tasks was indicated by a horizontal line that was continuously present in the middle of the screen. Specifically, stimuli presented at one or three positions above the line required a response based on color, and stimuli presented below the line required a response based on shape, or vice versa. In a task-cueing paradigm, each stimulus is preceded or accompanied by a cue, which serves to inform the participants which tasks has to be performed on the upcoming or current trial (e.g., Meiran, 1996; Sudevan & Taylor, 1987). In the current paradigm, a task cue preceded the target stimulus with an interval of either 150, 600, or 1500 ms that was fixed within trial blocks. The cue-to-target interval (CTI) was varied to assess developmental change in advance preparation. Given the results of Cepeda et al. (2001), advance preparation is reflected in decreased switch costs.

The task-switching paradigm was complicated by including go/nogo trials. Specifically, stimuli that occurred at the two most eccentric positions required a disjunctive response, rather than a choice response. In the present setup, one color (or shape) required response execution (go trial), whereas the other color (or shape) required response inhibition (nogo trial). A cue preceded the target stimulus to indicate the color (or shape) that signaled response execution (or inhibition). The CTIs were identical to those associated with choice response trials. Choice responses on trials, which are

preceded by a nogo trial, should be delayed relative to repeated choice responses (e.g., Hoffmann et al., 2003). Importantly, it is predicted that shifting from inhibition to responding takes disproportionally more time in children than adults (Band et al., 2000)[1]. At this point, it should be noted that the present experiment was not designed to assess prevailing notions of switch costs. Rather, the primary goal of the present study was to establish clearly that children's switch costs are disproportionally larger than adult's switch costs. This is compatible with either of the rivaling interpretations of the origin of switch costs. A second major aim of the present study was to assess children's flexibility in shifting from non-responding to responding. There is a rapidly growing literature on children's inability to shift from going to stopping (for a review, Van der Molen, 2000). However data pertaining to developmental differences in shifting from stopping to going are, to the best of our knowledge, lacking.

4.2   Method
4.2.1   Participants

Three groups of participants took part in this experiment: One group of 17 young children (7 female) ranging in age from 6 to 7 years old ($M = 7.33$ years, $SD = .88$) with IQs in the normal range (mean IQ = 111; $SD = 15$), a second group of 17 older children (8 female) ranging in age from 11 to 12 years old ($M = 11.80$; $SD = .62$) with IQs in the normal range (mean IQ = 106; $SD = 17$), and a third group of 15 young adults (10 female) ranging in age from 20 to 21 years old ($M = 20.40$; $SD = 1.64$) with normal IQs (mean IQ = 104; $SD = 9$; no IQ information was obtained from three participants because of software malfunction during IQ testing). Mean IQ did not differ between age groups, F $(2, 46) = 1.29$, $p = .286$. The children were recruited from local schools with the help of their teachers. Permission for the children to participate was procured from the children's primary caretakers. The young adults were recruited from a local university via flyers. The children received a small present for their participation, and the young adults received course credits. All participants had normal, or corrected-to-normal vision.

4.2.2   Apparatus and stimuli

An IBM compatible PC (486 processor, color monitor, SVGA graphics) was used for data acquisition and data analysis. On a light-gray background, a black horizontal line was continuously presented through the center of the screen. Perpendicular to this line, six squares of 2 x 2 cm were presented, with three squares above and three squares below the line. The distance between squares was 0.8 cm. Target stimuli were presented in these

---

[1] In the present study, "switching" refers to transitions from choice responses to shape to color (or vice versa), and "shifting" refers to transitions from choice to disjunctive responses (or vice versa).

squares and cue stimuli were presented adjacent to the squares, with 0.8 cm between cue and square.

Target stimuli were a blue circle, a yellow circle, a blue triangle, or a yellow triangle. The size of the target stimuli was approximately 1.3 x 1.3 cm. Cue stimuli differed in shape or in color. The cues differing in shape were a red circle and a red triangle, one presented at the left and the other at the right of the square in which the target would appear. The cues differing in color were a blue and a yellow square. On some trials, a black "X" was placed on one of the cue stimuli.

4.2.3 Task and trial sequences

A schematic of the display and the travel of the stimuli over the successive trials are shown in Figure 1. Target stimuli required either a left- or right-hand response (choice response), or the execution or inhibition of a response (disjunctive response). The "z" key of the computer keyboard was used for left index-finger responses, and the "/" key for right-index finger responses. The choice task was cued by the stimuli, which had the same color, but a different shape (i.e., red circle and red triangle), or which had the same shape, but a different color (i.e., yellow triangle and blue triangle). The same-color/different-shape stimuli cued the "respond-to-shape" choice task, and the same-shape/different-color stimuli cued the "respond-to-color" choice task. Thus, a red circle presented left to the square and a red triangle presented to the right indicated that the target stimulus required a left-hand response when it was a circle but a right-hand response when it was a triangle. Similarly, a yellow triangle presented at the left of the target square and a blue triangle at the right indicated that the target stimulus required a left-hand response, when it was yellow and a right-hand response when it was blue.

A black "X" placed on the stimuli, which appeared adjacent to the target square, cued the disjunctive task. Thus, crossing the red circle left to the target square, but not the red square appearing at the right, indicated that the left-hand response should be inhibited, if the target was a circle, but that the right-hand response should be executed if the target was a square. Finally, the position of the target provided an implicit task cue. Specifically, stimuli presented in the most eccentric squares required a disjunctive response, whereas targets presented in the four inner squares required a choice response. Targets presented above the horizontal line required responses based on the color of the target, whereas targets presented below the line required responses based on shape, or vice versa.
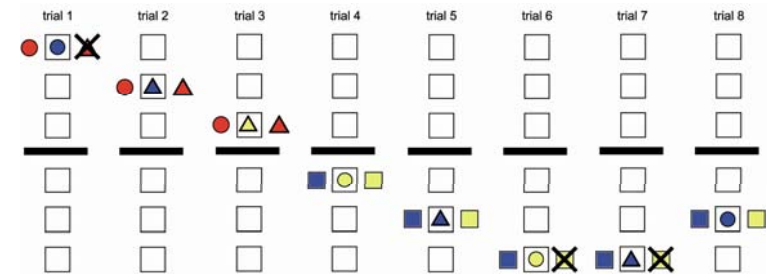
Figure 1. Schematic presentation of a trial sequence. The figure displays a sequence of the first eight trials. The leftmost display reflects the first trial, which required shape discrimination. The flanking triangle and the circle served as cues that instructed the participant to respond by pressing the left key ("z") when the target was a circle, and pressing the right key ("/") when the target was a triangle. Trial 2 and 3 also require shape discrimination. Trial 4 requires a switch to color discrimination, and the participant should press the left key for a blue target and the right key for a yellow target. Trial 5, 6, 7, and 8 also require shape discrimination, until trial 10 is reached and a shift to shape discrimination is required, etc. The top and the bottom array of the display were characterized by a disjunctive task. That is, a cue indicated whether a response should be made (Go) or withheld (NoGo). On these trials the square was flanked by a cue through which a cross was drawn, which indicated that no response should be made for that particular task

The cue-to-target interval (CTI) was 150 ms, 600 ms, or 1500 ms, and was varied between blocks of trials. There was a fixed delay of 1000 ms between the start of a trial block and the first cue. The first target and cue stimuli were always presented in the uppermost square (target), or adjacent to this square (cue). The next target was presented in the square just below the uppermost square, and the cue adjacent to it. The lowermost square was occupied twice and then the target traveled through the squares to the uppermost square that was then occupied twice.

4.2.4 Design and procedure

Each participant completed one test-session, consisting of 3 blocks of 78 trials. Across participants, the proportion of choice trials, go trials, and nogo trials was 67%, 16%, 17%, respectively. The mapping of stimuli onto responses were counterbalanced across participants and kept fixed across the experiment.

Test blocks were separated by a short rest of 2 min, and a longer rest of 5 min following the second test block. All participants first received a practice block consisting of 60 trials. They were instructed to balance speed and accuracy. Care was taken to ensure that participants understood the instructions. The task was administered in a dimly lit and sound-shielded room. Following the reaction time (RT) task, participants completed a computerized version of the Raven's Standard Progressive Matrices (Raven, Court, & Raven, 1985) as a measure of IQ. The duration of the total test session was approximately one hour.

4.3   Results

Prior to all analyses, RTs were transformed to their natural logarithm. This was done for two purposes: to minimize the influence of outliers (Ratcliff, 1993), and to reduce the influence of differences between age groups in baseline performance (Meiran, 1996). In this vein, switch costs are expressed in terms of ratio scores instead of absolute differences. Hence, an interaction between age group and an experimental variable will indicate a disproportional difference in RT between age groups in one condition relative to another condition. The results are reported in terms of the antilogs of the mean $nlog_{RTs}$ (i.e., the geometric means).

The analyses focused on several trial types that were defined by different trial sequences. The sequences consist of two trials: the current trial and the immediately preceding trial. Four sequences were considered that were defined by a choice response on the current trial. These choice responses could be preceded by a trial requiring the same choice response (i.e., choice-choice repeat trials, CCR), or by a trial requiring the alternative choice response (i.e, choice-choice switch trials, CCS). In addition, choice responses could be preceded by a disjunctive go response (i.e., go-choice shift trials, GCS), or by a disjunctive nogo response (i.e., nogo-choice shift trials, NGCS). Finally, three sequences were considered that consisted of a disjunctive go response on the current trial. These go responses could be preceded by another go response (i.e., go-go repeat trials, GGR), by a nogo response (i.e., nogo-go shifts, NGGS), or by a choice response (i.e., choice-go shifts, CGS).

The difference between CCS – CCR provides an estimate of switch costs similar to those obtained in the task-switching literature. The differences between GCS – CCR and between NGGS – GGR provide an estimate of the shift costs involved in switching from stopping to going. Finally, the differences between GCS – CCR and CGS – GGR provide estimates of the costs in switching from go to choice responses, and in shifting from choice to go responses, respectively.

Response Accuracy.   Three types of errors were possible: pressing the wrong key when the target requires a choice response (i.e., choice response errors), pressing the

wrong key on go trials (go errors), or pressing any key on nogo trials (nogo errors). All participants performed the task rather well. The proportion of choice response errors remained, on average, below 10% (1% in young adults, 1% in the older children, and 7% in the youngest children). The proportion of go errors was only slightly higher (4% in young adults, 2% in older children, and 10% in the youngest children). The proportion of nogo errors was highest (5% in young adults, 10% in older children, and 15% in the youngest children). Given the low error rates, these data were not analyzed in further detail.

Response Speed. Table 1 shows the RTs for the seven trial types that were designed to assess developmental differences in switch and shift costs. The first four responses in each block, responses that were preceded by errors, and trials with latencies shorter than 200 ms were excluded from analysis.

Trial type	young adults	11-12-year-olds	6-7-year-olds
Choice to *Choice* Repeat	431	580	848
Choice to *Choice* Switch	458	671	991
Go to *Choice* Shift	433	617	944
NoGo to *Choice* Shift	453	614	970
Go to *Go* Repeat	451	612	900
Choice to *Go* Shift	439	603	869
NoGo to *Go* Shift	416	601	967

Table 1. Geometric mean RTs (ms) as a function of trial type for each age group. Trial type refers to sequences of two trials; the current trial (in italics) and the immediately preceding trial. Response latencies are presented for the current trial. Repeat refers to repetition sequences. "Switch" to alternating choice responses (color vs. shape). "Shift" refers to transitions from choice to disjunctive responses, and vice versa

The ANOVA performed on choice responses included the four trial types CCR, CCS, GCS, and NGCS as within-subjects factor. The ANOVA indicated that response speed improved with advancing age. Mean RTs decreased from 937 ms in the 6-7 year olds, to 620 ms in the 11-12 year olds, and to 444 ms in the young adults, $F(2,46) = 84.93$, $p = .000$, $MSE = .3174$. When CTI was increased from 150 ms to 600 ms, mean RTs decreased from 679 ms to 600 ms. When CTI was increased to 1500 ms, the mean RT increased to 633, $F(2,92) = 22.23$, $p < .001$, $MSE = .0338$. Most importantly, there was a significant main effect of Trial Type, $F(3,138) = 27.02$, $p < .001$, $MSE = .0136$. Mean RT for choice repetitions was 597 ms, for choice alternations 673 ms. Mean RT for disjunctive go – choice shifts was 632 ms, and for disjunctive nogo – choice shifts it was

646 ms. These results show that both switching between choice responses to color vs. shape and shifting from disjunctive to choice responses, or vice versa, are both time consuming processes.

The ANOVA failed to show an interaction that included CTI but, of most importance, the interaction between Age Group and Trial Type did reach significance, F (6,138) = 3.05, p =.008, MSE = .0136. To further examine this interaction, follow-up analyses were done to evaluate age-related differences in switch and shift costs. Table 2 presents the costs involved in switching from one choice task to another and the costs involved in shifting from choice responses to disjunctive go responses, or vice versa. It can be seen that all costs are significantly larger in 6-7 year olds compared to young adults (planned contrasts, ps < .018). Switch costs, but not shifts costs, are significantly larger for 11-12 year-olds compared to young adults (p < .001).

Sequential costs	young adults	11-12-year-olds	6-7-year-olds
Choice to *Choice* Switch	27	91	143
	(1 vs. 2; 1 vs. 3)	(2 vs. 1)	(3 vs. 1)
Go to *Choice* Shift	2	36	95
	(1 vs. 3)		(3 vs. 1)
NoGo to *Choice* Shift	22	34	121
	(1 vs. 3)	(2 vs. 3)	(3 vs. 1; 3 vs. 2)
Choice to *Go* Shift	-12	-9	-31
NoGo to *Go* Shift	-34	-10	67
	(1 vs. 3)		(3 vs. 1)

Table 2. Switch and shift costs (in mean ms) as a function of trial type for each age group. Trial type refers to sequences of two trials; the current trial (in italics) and the immediately preceding trial. Switch costs are presented for the current trial relative to choice-to-choice repetitions. Shift costs are presented for the current trial (in italics) relative to choice-to-choice repetitions for choice responses and go-to-go repetitions for disjunctive responses. The numbers between parentheses refer to significant differences between particular age groups (1 = young adults, 2 = 11-12 year olds, and 3 = 6-7 year olds)

The ANOVA performed on disjunctive go responses included three trial types (GGR, CGS, and NGGS) as within-subjects factors. Mean RT decreased from 911 ms in the 6-7 year olds to 605 ms in 11-12 year olds, and to 435 ms in the young adults, F (2,46) = 102.02, p <.001, MSE = .1940. The main effect of CTI was significant, F (2, 92) = 33.39, p <.001, MSE = .0585. Increases in CTI from 150 ms to 600 ms to 1500 ms, were associated with decreases in mean RTs to 708 ms, to 598 ms, and to 567 ms, respectively. The main effect of Trial Type did not reach significance. The ANOVA failed to show an interaction that included CTI, but, importantly, there was an interaction of Type of Trial and Age Group, F (4,92) = 2.64, p =.039, MSE = .0386. In Table 2, it can be seen that in terms of mean RTs, NGGS were significantly larger in 6-7 year olds compared to young adults. Further analyses revealed that the go responses following nogo responses resulted in significant shift costs for the 6-7 year-olds, but not for the two other age groups, F (2,46) = 3.06, p < .35. Shifting from choice to go responses was without costs.

4.4   Discussion

This study set out to assess developmental differences in switch costs between tasks, and shift costs between go and nogo responses. Let us consider the adult findings first. Consistent with the task-switching literature, choice responses were reliably slower on switch trails compared to the responses on repeat trials (Allport et al., 1994; Meiran, 1996; Rogers & Monsell, 1995). Although there is a considerable amount of evidence that switch costs are reduced by an increase in the time available to prepare for a change in task prior to target onset (Meiran, 1996; Rogers & Monsell, 1995), the current findings failed to show such a beneficial effect of lengthening the cue-target interval. An explanation of the absence of a preparation effect is that the current manipulation of cue-target interval was not effective. However, this explanation is unlikely considering the robust effect of preparation time on the speed of responding. Thus, the latency of choice responses decreased about 80 ms when preparation time was increased from 150 ms to 600 ms. An alternative interpretation is that the absence of an appreciable effect of preparation on switch costs is due to the mixing of going and stopping in the current paradigm. Previously, Schuch and Koch (2003) observed only small, and sometimes unreliable, effects of preparation interval on switch costs when using a hybrid paradigm, which combined the switching and stopping of tasks. Subsequently, Kleinsorge and Gajewski (2004) conjectured, and indeed demonstrated, that individuals refrain from preparing when stop-trials are inserted in the series. Accepting this explanation for the current lack of a preparation effect, the current switch costs must be interpreted as residual; that is, a delay resulting either from interference (Allport et al., 1994; Wylie &

Allport, 2000), or from the insertion of an extra control process (Monsell et al., 2000; Rogers & Monsell, 1995).

The current findings are consistent with the literature on the aftereffects of stopping. For young adults, the delay in responding on choice trials following a nogo trial was about the same (22 ms) compared to the costs involved in switching choice tasks (27 ms). The delay of 22 ms is similar to those that have been observed in stop-signal studies (e.g., Kramer, Hahn, & Gopher, 1999; Rieger & Gauggel, 1999) and in task-switching studies (e.g., Hoffman et al., 2003; Kleinsorge & Gajewski, 2004; but see Schuch & Koch, 2003, for stronger effects). The after-effect in stop-signal trials has been interpreted in terms of a specific inhibitory mechanism similar to negative priming (e.g., Rieger & Gauggel, 1999), whereas the costs of shifting from a stopping to going in the task-switching paradigm has been interpreted in terms of a response repetition benefit operating on going-going trials (Hoffman et al., 2003). The latter interpretation is less likely, however, in view of the substantial literature on sequential effects showing a relatively rapid decay of repetition benefits—usually within 500 ms for adult subjects (Kirby, 1980; Smulders et al., 2004; Soetens et al., 1985). An alternative interpretation involves a hierarchical control model in which an executive mechanism issues commands to subordinate processes in a response-production system (e.g., Logan & Cowan, 1984; Meyer & Kieras, 1997a, 1997b; Norman & Shallice, 1986; Stuss et al., 1995). On the decision to stop an activated motor response, the executive issues a command to the subordinate system that, if successful, results in the silencing of the motor system. When the subsequent trial requires a speeded motor response, the executive issues a command to the subordinate system to release the inhibitory state, and to activate and execute the designated response. The current finding of nogo-to-choice shift costs suggests that releasing the inhibitory state (i.e., the inhibition of inhibition) is a time-consuming process.

The disjunctive responses observed for adult participants showed a somewhat surprising pattern, with longer latencies associated with go-go repetitions compared to nogo-go shifts, and approximately equal latencies compared with shifts from choice to go responses. The finding that stopping has aftereffects (Rieger & Gauggel, 1999), and the observation that choice responses are delayed following a nogo trial (present study, Kleinsorge & Gajewski, 2004; Schuch & Koch, 2003) suggested that go responses following go responses would be delayed relative to go-go repetitions. But the results showed the opposite result. One interpretation of this unexpected finding refers to the sequential effects literature (e.g., Kirby, 1980; Luce, 1986; Soetens et al., 1985). This literature shows that choice repetitions are slower than choice alternations when the response-to-stimulus interval is longer than 500 ms. In the present study, the interval between trials was considerably longer. Thus, the sequential effect observed previously for choice responses may apply also to disjunctive responses. More specifically, the loss

in time resulting from the need to shift from stopping (on nogo trials) to going (on go trials) is compensated by the beneficial effect of an alternating sequence. In the sequential effects literature, the beneficial alternation effect is usually interpreted in terms of subjective expectancy, assuming that participants tend to overestimate the probability that an alternation will occur, when actually the probabilities of repetitions and alternations are equal (e.g., Smulders et al., 2004; Soetens et al., 1985; Sommer, Leuthold, & Soetens, 1999).

Finally, the speed of go responses following choice responses was about the same as the speed of choice responses following go responses (i.e., 439 ms for choice-go shifts and 433 ms for go-choice shifts). This finding is remarkable given that, when presented in blocks of trials, go responses are typically much faster than choice responses (e.g., Luce, 1986; Van der Molen & Keuss, 1981; Woodworth & Schlosberg, 1954). Obviously, it should be acknowledged that the probability of a choice response was twice the probability of a go response. The other finding of interest is that shifting from choice to go responses and vice versa does not seem to involve shift costs, when choice repetitions (431 ms) are used as a baseline. This finding is of interest as it may provide indirect support for Donders' (1868) subtraction logic. Donders assumed that choice responses and disjunctive go responses are identical, with the exception of a response-selection stage that is involved in choice responses, but absent in disjunctive go responses. Accordingly, subtracting disjunctive RT from choice RT would provide an estimate of the duration of the response selection stage. When interpreting switch costs as reflecting interference at the response selection stage (Allport et al., 1994; Meiran et al. 2000), the current lack of go-choice shift costs may suggest that disjunctive go responses do not entail a selection stage (i.e., the selection between a nogo vs. go response).

Let us consider now the developmental differences that emerged from the current experiment. The results showed a clear-cut developmental decrease in switch costs; from a mean of 163 ms in 6-7 year olds, to a mean of 89 ms in 11-12 year olds, and a mean of 27 ms in young adults. This result indicates that the ability to switch from one task to another continues to develop well beyond childhood. The current finding of a developmental increase in task-switching ability is consistent with previous studies that examined developmental change in cognitive flexibility using the task-switching paradigm (Cepeda et al., 2001; Crone et al., 2006). The increasing ability to switch tasks has been interpreted vis-à-vis the two dominant interpretations of switch costs. Thus, it has been argued that the larger switch costs in children reflect immature levels of executive control (Cepeda et al., 2001; Zelazo et al., 2004). By contrast, Crone et al. (2006) interpreted larger switch costs in children in terms of memory retrieval. More specifically, this account posits that, when a response is linked to particular stimulus on one trial, the S-R link is encoded in memory. If the stimulus occurs again on the next trial, a response tag is automatically retrieved. This response tag will then interfere with

the current response requirement. Crone argued that children experience greater difficulty in resolving the response interference. The current findings, however, cannot discriminate between the two alternative interpretations[2].

The current findings showed also a pronounced developmental trend in the costs associated with the shift from stopping to choice responses. Interestingly, the time course of these shift costs was different from the time course of switching costs. The ability to switch tasks continued to develop into adolescence whereas the ability to shift from stopping to going seems to develop more rapidly—the current results showed little difference between 11-12 year olds and young adults in this regard. The developmental change in the ability to shift from stopping to going is consistent with the previous findings reported by Band et al. (2000), who showed that young children are disproportionally slower than young adults in shifting to the alternate response in a stop-change task. Within a hierarchical control model (Logan & Cowan, 1984; Meyer & Kieras, 1997a, 1997b; Norman & Shallice, 1986; Stuss et al., 1995) children's inefficient way in handling stop-go sequences is interpreted in terms of their difficulty in relinquishing the inhibitory state induced on the previous nogo trial. This interpretation is supported by the additional finding that go responses following a nogo trial are considerably slower than go responses following a nogo trial, at least for the youngest children. This latter finding is opposite to the results observed for young adults. The current interpretation of the results for adults assumed that the time consumed by the need to release the inhibitory state was compensated by their expectancy of an alternation to occur (i.e., a go trial). Why is this compensation absent in children? The answer is simple—several studies indicated that subjective expectancy is absent in children (e.g., Kerr, Blanchard, & Miller, 1980; Kerr, Davidson, Nelson, & Haley, 1982; Smulders et al., 2004; Soetens & Hueting, 1992).

In conclusion, the results that emerged from the present study revealed two clear developmental trends: (i) the ability to switch choice responses continued to develop (i.e., improve) into adolescence, and (ii) the ability to shift from stopping to going attained mature levels in late childhood. The former trend is consistent with a several previous studies that examined task-switching in children. The latter trend is a new contribution to the rapidly growing literature on developmental change in the ability to inhibit. Both trends can be conceptualized in terms of a hierarchical model of control. Within this framework, switching and shifting are interpreted in terms of an extra control process delaying the speed of responding on a particular trial (e.g., Monsell et al., 2000; Rogers &

---

[2] Fine-grained analyses of trial sequences may shed light on this issue (e.g., by focusing on longer sequences (e.g., Mayr & Keele, 2000; Schuch & Koch, 2003), or on sequences sharing stimuli or responses (e.g., Crone et al., 2004; Meiran et al., 2000). Unfortunately, the current number of trials is too small for performing more detailed analyses.

Monsell, 1995; see also Meyer & Kieras, 1997a, 1997b). The current contribution is twofold. First, the present findings indicate that the control mechanism implicated in switching between choice responses is slow to mature. Second, the present findings demonstrate that releasing inhibition is a time-consuming process. In addition, the ability to release inhibition seems to attain adult levels during late childhood. This pattern is consistent with the developmental trends observed previously by Band et al. (2000). These authors observed that the ability to inhibit developed more rapidly than did the ability to recruit and execute a choice response.

It should be noted, however, that the current results are also compatible with a model assuming competition or interference among subordinate processes (Allport et al., 1994; see also Logan & Bundesen, 2003). Within such a framework, it is assumed that individuals employ exactly the same set of processes as on the previous trial. However, at least one process, presumably response selection, is prolonged due to interference effects from the previous trials. Accordingly, developmental differences in switching costs are then interpreted in terms of children's greater sensitivity to interference or their inefficient ways of handling such interference. Ridderinkhof and Van der Molen (1995) obtained strong evidence for a developmental trend in the resistance to interference. They had children in different age groups perform on an Eriksen flanker task, and observed that the costs in responding to a central target, which was surrounded by flankers associated with the alternate response, were considerably larger in childhood, compared to those observed in adolescence and adulthood. Moreover, they observed that the costs were not due to perceptual conflict, but rather to competition at the level of response activation, as indicated by brain potential measures of stimulus evaluation (P3) and preferential response activation (Lateralized Readiness Potential, LRP). It would be of considerable interest to use the same brain potential methodology to assess the locus of developmental trends in the ability to inhibit inhibition (i.e., the ability to release inhibition).

Chapter 5 ● Age-related change in shifting attention between global and local levels of hierarchical stimuli

Abstract

The current study examined the developmental pattern of the ability to shift attention between global and local levels of hierarchical stimuli. Three age groups (i.e., 7-year-olds, 11-year-olds, and 21-year-olds) performed on a task that allowed for the examination of 1) the direction of attention to either global or local levels of a stimulus (i.e., global or local preference); 2) the susceptibility to interference of the global or local levels of hierarchical stimuli; and 3) the flexibility in directing attention to global or local levels of hierarchical stimuli. Two experiments were carried out. In the first experiment the shape of a cue drew attention to the global or local level of the target stimulus. In the second experiment the target level was indicated by symbolic cues. The results indicated a global precedence effect that decreased with advancing age when the task required level shifting from trial-to-trial. The abilities to resist interference and to flexibly shift between global-local levels also improved during childhood, but quickly leveled off during adolescence. Cue type did not interact with these developmental trends. These findings are important vis-à-vis the orthogenetic principle, which implies a developmental change towards local analysis. The local-global paradigm was found to be a useful tool for examining the developmental course of multiple attention components.

5.1   Introduction

Attention involves the preparedness for, and selection of, sensory input from the physical environment (Raz & Buhle, 2006), and plays a critical role in the normal human cognitive, emotional, and social development (e.g., Johnson, 1998). The nature of attention is complex, and previous studies suggested multiple components of attention. For example, the influential model of attention of Posner and Petersen (1990) involves three functionally and anatomically distinct functions (networks) of attention: alerting (i.e., the ability to increase and maintain response readiness in preparation for an imminent stimulus), orienting (i.e., the ability to select specific information from among multiple sensory stimuli), and executive attention (i.e., the ability to monitor and resolve conflict between computations in different neural areas; for a review see Raz & Buhle, 2006). Recently, Fan, McCandliss, Sommer, Raz, and Posner (2002) developed the attention network test (ANT) to provide a behavioral measure of the three attentional

networks within one single task. The ANT is based on the flanker task (Eriksen & Eriksen, 1974) and the cued RT task (Posner & Petersen, 1990), and requires the participant to determine whether the middle arrow of five vertically arranged arrows points left or right. Different cueing conditions (i.e., alerting cues, spatial valid and invalid cues, and congruent and incongruent flankers) provide an index of the efficiency of the different attentional networks (Fan et al., 2002). Recent neuroimaging studies provided support for the anatomical distinctness of the proposed attentional networks (for a review see Raz & Buhle, 2006).

Only recently, the ANT has been used to assess the developmental pattern of attentional mechanisms (Konrad et al., 2005; Rueda et al., 2004). For example, Konrad et al. (2005) used structural and event-related functional imaging to asses the development of alerting, orienting, and executive attention functions in 8 to 12-year-old children, and in adults. At the behavioral level, children were impaired on the three attentional mechanisms. At the brain level, during development, a shift from functional yet immature systems underlying attentional functions to more focal adult networks was observed. Konrad et al. (2005) interpreted the results as reflecting both developmental changes in cognitive strategies and brain maturation.

The current study shared the goal of the ANT studies of attention development by focusing on the developmental trajectories of distinct attention components. However, it employed the 'global-local' paradigm, which allows for the assessment of complementary attention components; i.e., attention focus, sensitivity to interference, and flexibility. The global-local paradigm was developed to study the ability to integrate the parts and the whole of visual scenes (Navon, 1977). This experimental paradigm makes use of hierarchically organized stimuli with two interrelated levels, the local and the global level, where the latter level (the whole) comprises elements arranged at the former level (the parts). Studies that addressed global-local processing in adults typically used hierarchical (compound) stimuli, e.g., a large *H* consisting of smaller letters that are either congruent (e.g., small *H*s) or incongruent (e.g., small *S*s; see Navon, 1977). The participant is required to identify the stimulus at either the global level or the local level. To this end the task includes two types of trials, congruent and incongruent trials. On congruent trials the critical information at global and local levels of the hierarchical stimulus is mapped onto the same response, while on incongruent trials the information at global and local levels activates competing responses. Thus, the global-local paradigm allows for the examination of three mechanisms of attention: 1) the direction of attention to either global or local levels of a stimulus (i.e., global or local preference; Navon, 1977); 2) the susceptibility to interference (i.e., the ability to resist the interference of irrelevant stimulus information; Eriksen & Eriksen, 1974); and 3) the flexibility of attention (i.e., the ability to flexibly shift the direction of attention between target levels; e.g., Shedden, Marsman, Paul, & Nelson, 2003). The three attention components that can

be examined using the global-local paradigm will be reviewed briefly below, followed by a set of predictions concerning the developmental trajectories of these attention components.

Global or local precedence. Using the global-local paradigm, Navon, in his seminal work (1977), obtained a "global precedence" effect, showing that participants identified the information at the global level of a stimulus significantly faster than they do at the local level. This effect has been replicated in numerous studies (for reviews see Kimchi, 1992; Navon, 2003, see also Miller & Navon, 2002). Subsequent studies indicated, however, that the results obtained by Navon (1977) and others are critically dependent on task characteristics. That is, several studies reported a "local precedence" effect by manipulating stimulus features, such as increasing the overall stimulus size (Enns & Kingstone, 1995; Kimchi, 1992; Kinchla & Wolfe, 1979; Lamb & Robertson, 1990; McLean, 1979; Navon, 1983), by decreasing the number of local elements (Kimchi, 1988; Kimchi, Hadad, Behrmann, & Palmer, 2005; LaGasse, 1993; Martin, 1979), or by presenting the stimulus in the left or right hemi-field (e.g., Delis, Robertson, & Efron, 1986; Kimchi & Merhav, 1991; Lamb, Robertson, & Knight, 1990; Robertson, Lamb, & Knight, 1991). Nonetheless, across a wide range of conditions that intended to eliminate most confounds, the prevalent finding is that the global level is perceived first (e.g., Enns, Burack, Iarocci, & Randolph, 2000; Goto, Wills, & Lea, 2004; Miller & Navon, 2002; Navon, 2003).

A central theme in developmental studies of the ability to integrate the parts and the whole of visual scene has been framed by the "orthogenetic principle" (Werner, 1948, 1957). This principle states that development of perception of stimuli with multiple levels proceeds from an initially undifferentiated state to one increasing specialization, and finally to the coordinated integration of specialized components (cf. Enns et al., 2000). That is, when children grow older, the development of perception follows a global-to-local pathway. This notion can be illustrated by face perception: young children first perceive the spatial contiguity of a face, rather than its constituents. As they grow older, the perception shifts towards the constituents (for a review see Maurer, Le Grand, & Mondloch, 2002). This global-to-local development is due to the protracted development of the visual system, which ultimately gives rise to the ability to perceive local features and its spatial relations (i.e., the global structure; Enns et al., 2000; Maurer et al., 2002). Thus, given the apparent global precedence effect found in adults, and the global-to-local developmental pathway in children (i.e., the orthogenetic principle) on the other hand, one would expect a greater global precedence effect in young children. The outcomes of studies of the developmental pattern of global-local processing are rather contradictory. However, consistent with the results that emerged from adult studies, the age-related patterns seem to be highly dependent on task characteristics, such as different stimuli and displays (Enns et al., 2000; Kimchi et al., 2005).

In the present study we studied developmental change (during childhood, and into young-adulthood) in global-local processing by using stimuli that have been used frequently studied in adults, and have yielded clear global precedence effects (Kimchi, 1992; Kimchi & Palmer, 1985). The spatial pattern of these stimuli consisted of many, relatively small, elements (see Figure 1). The participants performed on two tasks—the global and local tasks. In the global task, they were instructed to respond to the large global shape of the target stimulus (i.e., a square or a rectangle). In the local task, they were instructed to respond to the shape of small elements that comprised the target stimulus (i.e., squares or rectangles). A comparison of performance on the "respond-to-global" trials and the "respond-to-local" trials yielded the global precedence index. Based on previous reports using the current stimuli (Kimchi, 1992; Kimchi & Palmer, 1985), it was anticipated that adults would show global precedence. More importantly, based on the orthogenetic principle (Enns et al., 2000), it was predicted that global precedence would be strongest in the youngest children, and that this precedence would decrease with advancing age.

Susceptibility to interference. Hierarchical target stimuli, as used in the global-local paradigm, include a relevant aspect (i.e., the current task to perform, e.g., "respond-to-global"), and an irrelevant, distracting, aspect (i.e., the task to ignore; e.g., "respond-to-local"). As such, target stimuli can either be congruent (i.e., the relevant and the irrelevant level require the same response), or incongruent (i.e., the relevant and the irrelevant level require a different response). On incongruent trials, interference of global information in the analysis of local information, or vice versa, occurs. The ANT studies compared performance on congruent trials and incongruent trials to provide an index for the ability to resist interference. The typical finding is that incongruence is associated with slower and less accurate responses, even though participants are instructed to ignore the distracting aspect of a stimulus (e.g., Fan et al., 2002; Raz & Buhle, 2006; see also Eriksen & Eriksen, 1974). Studies using the global-local paradigm consistently indicated that irrelevant global information disrupts task-relevant analysis of local information (i.e., global-local interference), whereas the opposite does not occur (i.e., asymmetric interference; e.g., Navon, 1977; see also Kimchi, 1992; Navon, 2003). These findings led us to predict that in adults an asymmetric interference effect will be observed. In addition, we predicted a developmental decrease in the effect of interference. More specifically, global-local interference will be most pronounced in young children, whereas the developmental trend for local-global interference will be less pronounced (see also Konrad et al., 2005; Rueda et al., 2004).

Flexibly directing attention. The global-local paradigm can be used also to examine the ability to flexibly shift the direction of attention between global and local levels of a target stimulus. More specifically, this ability can be examined by presenting the participants with hierarchical stimuli on two consecutive occasions, and determining

the participants' speed in identifying the target stimulus on the second stimulus presentation (e.g., Hübner, 2000; Shedden et al., 2003; Ward, 1982). Two sequences can be considered: (1) the target stimulus remains at the same global-local level (e.g., the participant has to identify the target stimulus at the global level on the first trial and at the same level on the second trial; i.e., level repetition); and (2) the target level changes on the second trial (e.g., the participant has to identify the target stimulus at the global level on the current trial, and on the local level on the subsequent trial; i.e., level alternation). Previous studies in adults, requiring the global-local analysis of hierarchical stimuli, found that responses to the second trial were executed faster on level repetition trials than on level alternation trials (i.e., the level-shifting effect; Filoteo, Friedrich, & Stricker, 2001; Hübner, 1997, 2000; Lamb & Yund, 1996, 2000; Robertson, 1996; Shedden et al., 2003; Ward, 1982). Importantly, most of these studies demonstrated that the level-shifting effect occurred regardless of whether there was a bias towards processing or local levels of the stimuli, indicating that the cognitive mechanisms involved in global or local precedence may be distinct from those mechanisms involved in the level-shifting effect (e.g., Lamb & Yund, 1996; Robertson, 1996; Ward, 1982). Based on this literature, we predicted that in adult participants alternating between target levels would occur with a higher cost compared to responding to repeating target levels, regardless of the hierarchical level of the target stimulus that was to be shifted to. In addition, it was anticipated that the costs involved in level-shifting would be most pronounced in young children. This prediction was based on recent developmental studies of task-switching in which participants were required to switch between two rules of behavior. These studies showed that the ability to switch between two tasks developed gradually during childhood, and reached adult levels of performance at around age 12 (e.g., Cepeda, Kramer, & Gonzalez de Sather, 2001; Crone, Bunge, Van der Molen, & Ridderinkhof, 2006; Huizinga, Dolan, & Van der Molen, 2006, but see Kray, Eber, & Lindenberger, 2004). This hypothesis can be further specified by positing separate trajectories for each processing level, as reflected in the developmental findings that indicate that adult level of sensitivity is attained earlier for local than for global target levels (e.g., Burack, Enns, Iarocci, & Randolph, 2000; Kimchi et al., 2005). Moreover, given the hypothesis of an orthogenetic trend from global to local spatial processing, we expect that in children shifting attention from global to local levels will be associated with a higher cost compared to shifting from local to global levels.

5.1.1 The present study

In sum, the current experiment was designed to examine the development of 1) the direction of attention to either global or local levels of a stimulus (i.e., global or local preference); 2) the susceptibility to interference; and 3) the flexible direction of attention

to global or local levels of hierarchical stimuli. In a binary choice RT task participants of three age groups (7-, 11-, and 21-year-olds) were required to respond as fast as possible to the global or the local level of hierarchical stimuli, which were adapted from Kimchi (1992) and Kimchi & Palmer (1985). Half of the stimuli was congruent, and the other half was incongruent. In addition, we adopted from the task-switching literature a hybrid version of the "alternating-runs" and the "explicit task-cuing" procedures. The alternating-runs procedure consists of two tasks (e.g., "respond-to-global", and "respond-to-local"), and the task changes predictably every nth trial (e.g., Rogers & Monsell, 1995); in the explicit task-cuing procedure participants are first given a cue that indicates which task to perform, followed by a target stimulus on which to perform the cued task (e.g., Meiran, 1996; Sudevan & Taylor, 1987). We expected 1) a global precedence effect in each age group, being greatest in the youngest children and decreasing with advancing age; 2) a developmental decrease in the susceptibility to interference with larger interference from global than from local level (this asymmetry is anticipated to decrease when children grow older); and 3) an age-related decrease in the costs involved in level-shifting.

The current study involved two experiments. In the first experiment, participants were presented with task cues that are relatively easy to match to the target stimulus (i.e., the target stimuli could be spatially matched to the cue stimuli). In this experiment, the shape of the cue and the target-level of the hierarchical stimulus allowed for a physical matching of the cue and the target, thereby reducing the task demands such that young children would be able to perform the task. It should be acknowledged, however, that the opportunity of a physical match may reduce of the chances of observing potential developmental trends in global-local processing. Thus, a second experiment was conducted using symbolic cues for directing attention to global vs. local levels of the target stimulus. Moreover, the results of the second experiment allowed us to assess the consistency of developmental change in global-local analysis across the two types of cues.

5.2 Experiment I
5.2.1 Method
5.2.1.1 Participants

Three age groups (7-year-olds, 11-year-olds, 21-year-olds) were tested. Children were recruited from regular local schools; the 21-year-olds were university students. The children received a small present for their participation, and the 21-year-olds received course credits. Informed consent was obtained. All participants reported to be healthy, and had normal, or corrected-to-normal, vision. The Standard Progressive Matrices (SPM; J. C. Raven, Court, & Raven, 1985) was used to obtain an estimate of cognitive

level. Scores were converted to quartile scores, following the norms for each age group. There was no significant difference between age groups on Raven SPM quartile, F (2, 58) = 1.27, p = .287. In addition, chi-square analyses indicated that gender distribution did not differ significantly between age groups, χ^2 (2) = .132, p = .936. The distribution of age and gender, and Raven quartile scores are presented in Table 1.

Age Group	Age			N female (male)		Raven quartile	
	Exp 1	Exp 2		Exp 1	Exp 2	Exp 1	Exp 2
7-year-olds	7.22 (.32)	7.39 (.65)		10 (9)	15 (9)	3.6 (.49)	2.6 (1.24)
11-year-olds	11.21 (.43)	11.10 (.58)		11 (10)	11 (10)	3.5 (.51)	2.7 (.73)
21-year-olds	20.75 (2.20)	20.91 (2.38)		13 (9)	13 (9)	3.7 (.46)	2.3* (.66)

Table 1. Number of participants, mean age (years; SD in parentheses), gender, mean Raven SPM quartile (SD in parentheses) per Age Group in Experiments 1 and 2; * Raven APM percentile

5.2.1.2   Apparatus

Task administration was computerized. The tasks were presented on a Toshiba Satellite 1600 laptop (Intel Celeron 800 mHz processor; 15 inch 60 Hz monitor). All tasks required only left- and right-hand responses. The response button for the left hand was the "z" key on the computer keyboard, and for the right hand, the "?" key (responses were counterbalanced across participants).

5.2.1.3   Stimuli

The target stimuli were adopted from Kimchi (1992) and Kimchi and Palmer (1985), and consisted of geometric figures. Larger (global) rectangles/squares consisted of smaller (local) rectangles or squares. Global stimuli (i.e., squares or rectangles; 93 x 93 pixels or 93 x 189 pixels respectively) were composed of many smaller "local" stimuli (i.e., squares or rectangles; 21 x 21 pixels or 8 x 46 pixels respectively). The space between the local elements of a stimulus was 3 pixels. A global square consisted of 16 small squares or 8 small rectangles; a global rectangle consisted of 32 small squares or 16 small rectangles. A schematic of the stimuli is presented in Figure 1.
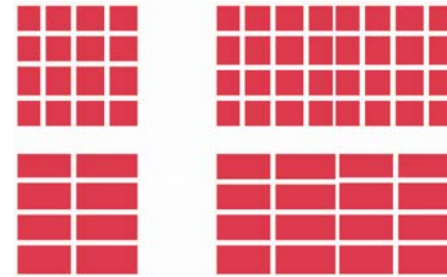
Figure 1. The target stimuli used in Experiments I and II. Clockwise, departing in the right-upper quadrant: big rectangle consisting of small squares (incongruent); big rectangle consisting of small rectangles (congruent); big square consisting of small rectangles (incongruent); big square consisting of small squares (congruent)

5.2.1.4   Procedure

A black horizontal line was presented continuously in the middle of the computer screen, against a light-gray background. The target stimuli were presented above or below the horizontal line. The task of the participant was to respond to the global or the local level of the target stimulus with either a left- or right-hand button press depending on the level indicated by the task cue stimulus. The mapping of the responses onto the stimuli was counterbalanced across participants – with the constraint that for a particular participant responding to (small or big) squares was mapped onto one hand, and responding to (small or big) rectangles was mapped onto the other hand. Half of the trials required a right-hand response; the other half required a right-hand response. In addition, the presentation of the respective tasks above or below the horizontal line was counterbalanced across participants (i.e., "attend-to-global" was presented above the horizontal line and "attend-to-local" was presented below the line, or vice versa).

The global-analysis condition was indicated by large cues presented above (or below) the horizontal line on which the target appeared, whereas the local-analysis condition was indicated by small cues below (or above) the horizontal line. The large cues consisted of a big square, which was presented at one side of the target stimulus, and a big rectangle presented at the other side of the target stimulus; the small cues consisted of a small square and a small rectangle. Cues were presented at 1 cm from the target stimulus. The size of global and local cues was the same as the target stimuli, but the global cues did not consist of smaller elements. The color of the target stimuli and the cues was red.

The time interval between cue onset and the presentation of the target stimulus was 500 ms. The cue and the target stimuli remained on the screen until a response was given. Participants had 3500 ms to respond. The response initiated the cue for the next trial with a fixed delay of 1000 ms. A schematic of the trial sequence is presented in Figure 2.
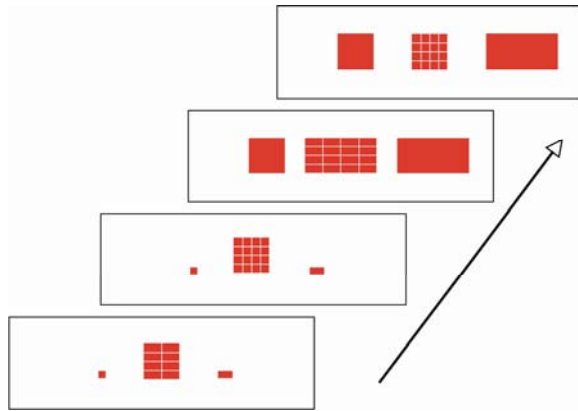

Figure 2. The figure displays a trial sequence proceeding from bottom to top, in a mixed-level block. The first display is the last event of the previous trial, which required local discrimination. The small square on the left and the rectangle on the right served as cues that instructed the participant to respond by pressing the left key when the target consisted of small squares and by pressing the right key when it consisted of small rectangles. After each response, a response-cue interval occurred, followed by the presentation of a cue, and after another interval the target appeared. The second display indicates a task repetition, the third display indicates a task shift to global discrimination. The big square on the left indicated that the participant should press the left key for a big square and the right key for a big rectangle

The participants underwent three blocks of trials, i.e., two "single-level" blocks, and one "mixed-level" block. In one block, the participants performed the local task, in a second block, they performed the global task (i.e., single-level blocks 1 and 2, in randomized order; 30 practice trials and 50 experimental trials per block), and in a third block, they alternated between a series of four repetitions of global-directed trials and a

series of four repetitions of local-directed trials (i.e., mixed-level block 3; 90 practice trials, 150 experimental trials).

The participants were coached to balance speed and accuracy in responding. Care was taken to ensure that participants understood the instructions, as indexed by verbal report, response accuracy, and stable RTs. The main dependent variables were the square roots of the error percentages and median RTs.

The test was administered in groups of two. At the end of the test-session, the 21-year-olds completed a paper-and-pencil version of the Raven SPM (Raven et al., 1985). The children completed this task in the classroom.

5.2.1.5 Design

In both the "single-level" blocks and the "mixed-levels" block, there were two target conditions: the "respond-to-global" and the "respond-to-local" condition. In addition, in both the "single-level" blocks and the "mixed-levels" block, interference was manipulated by the congruency of the stimuli. That is, target stimuli were either congruent (i.e., global and local levels shared the same shape, namely, square or rectangle) or incongruent (i.e., global and local levels differed in shape, square or rectangle). Moreover, in the mixed block, the focus was on two trial types that referred to two different trial sequences consisting of two trials—the current trial and the immediately preceding trial. That is, a choice reaction could either be preceded by a trial requiring a choice reaction either to the same task (e.g., "attend-to-global" preceded by "attend-to-global"; i.e., level repetition) or by a trial requiring a choice reaction to a different task (e.g., "attend-to-global" preceded by "attend-to-local"; i.e., level alternation).

5.2.2 Results

The results are presented in two sections. The first section focuses on the performance on the single-level blocks. This allows for the examination of developmental change in global precedence, and in the susceptibility to between-levels interference. The second section focuses on performance in mixed-level blocks, allowing for an examination of the development of the ability to flexibly shift between global and local levels of hierarchical stimuli, and to examine interference effects. In the analyses reported below, Age Group (7-year-olds, 11-year-olds, 21-year-olds) was included as between-subjects factor. The first five trials in each task block, trials with RTs shorter than 120 ms were excluded from the analyses. In addition, error trials and trials following an error were excluded from the RT analyses.

5.2.2.1   Single-level blocks

The square root of error percentages and median RTs were submitted to separate repeated measures analyses of variance (ANOVAs), with Target Level (global vs. local) and Congruency (congruent vs. incongruent) as within-subjects factors.

Errors. The ANOVA revealed a main effect of Target Level that reflected global precedence, as indexed by a smaller proportion of errors on global-directed trials compared to local-directed trials (1.8% vs. 2.8%), $F(1, 58) = 5.46$, $p = .023$. Moreover, a main effect of Congruency reflected interference of the irrelevant target level, as indicated by a smaller proportion of errors on congruent trials compared to incongruent trials (1.3% vs. 3.6%), $F(1, 58) = 24.19$, $p < .0001$. Finally, a significant interaction of Target Level and Congruency revealed that an asymmetric effect of interference, in that local-directed trials suffered more interference from the global level than global-directed trials did from the local level (3.9% vs. 1%), $F(1, 58) = 7.52$, $p = .008$. No significant main or interaction effects involving Age Group were found. Thus, any between-Age Group differences in response speed cannot be interpreted in terms of a trade-off between accuracy and speed.

Response latencies. The ANOVA revealed a main effect of Age Group, reflecting shorter response latencies in older children (753 ms in 7-year-olds, 504 ms in 11-year-olds, and 367 ms in 21-year-olds), $F(2, 58) = 134.78$, $p < .0001$. In addition, the significant main effect of Target Level reflected global precedence, as indicated by faster responses to global-directed trials compared to local-directed trials (503 vs. 579 ms), $F(1, 58) = 83.83$, $p < .0001$. Contrary to the expectations, however, there was no difference between Age Groups with respect to this effect, $F(2, 58) = 1.13$, $p = .331$. Furthermore, a significant main effect of Congruency reflected interference of the non-target level, as shown by faster responses to congruent trials compared to incongruent trials (528 vs. 555 ms), $F(1, 58) = 29.42$, $p < .0001$. Again, and contrary to the expectations, there was no difference between Age Groups with respect to this effect ($F < 1$). A significant interaction of Target Level and Congruency, $F(1, 58) = 4.98$, $p = .030$, indicated an asymmetric effect of interference - local-directed trials suffered greater interference than global-directed trials (39 ms vs. 17 ms). No further significant effects were found.

5.2.2.2   Mixed-level block

In addition to the examination of developmental change in global precedence and susceptibility to interference, the data from the mixed block allowed for an assessment of level-shifting. More specifically, two different trial sequences consisting of two trials were examined, namely level repetitions (the target level on trial N was the same as the target level on trial N-1), and level alternations (the target level on trial N differed from

the target level on trial N-1). The square root of the error percentages and median RTs were submitted to separate repeated measures ANOVAs with Target Level (global vs. local), Congruency (congruent vs. incongruent), and Level Type (level repetition vs. level alternation) as within-subjects factors.

Errors. The ANOVA revealed a significant main effect of Age Group, indicating a decrease in erroneous responding in older children (4.4% 7-year-olds, 2.8% in 11-year-olds, and 1.7% in 21-year-olds), $F(2, 58) = 5.45$, $p = .007$. Further, Target Level had a significant effect that reflected global precedence, as indicated by fewer errors on global-directed trials as compared to local-directed trials (1.2% vs. 3.6%), $F(1, 58) = 13.20$, $p = .001$. There was no difference between Age Groups with respect to this effect. Finally, the significant main effect of Congruency indicated interference from the non-target level, as reflected by fewer errors on congruent trials as compared to incongruent trials (0.8% vs. 6.1%), $F(1, 58) = 76.01$, $p < .0001$, but there was no interaction with Age Group. The main effect of Level Type was not significant, but there was a significant Age Group x Level Type interaction, $F(2, 58) = 3.57$, $p = .034$. Post-hoc analysis indicated that the effect of Level Type was significant only in adult participants ($p = .003$); in adults, a decreased accuracy of .8% on repetition trials was observed. All other effects failed to reach significance.

Response latencies. The ANOVA revealed a significant main effect of Age Group, indicating shorter response latencies with advancing age (894 ms in 7-year-olds, 556 ms in 11-year-olds, and 409 ms in 21-year-olds), $F(2, 58) = 121.82$, $p < .0001$. Furthermore, a significant effect of Target Level was obtained that indicated global precedence, as indexed by faster responses to global-directed trials compared to local-directed trials (569 vs. 670 ms), $F(1, 58) = 130.29$, $p < .0001$. As expected, a significant interaction of Age Group and Target Level showed that in all age groups the responses to global-directed trials were faster than the responses to local-directed trials. Importantly, this difference decreased with age (162 ms in 7-year-olds, 89 ms in 11-year-olds, and 50 ms in 21-year-olds), $F(2, 58) = 22.42$, $p < .0001$. Subsequent contrast analyses revealed that adult level of performance was attained in 11-year-olds.

A significant main effect of Congruency was also obtained, $F(1, 58) = 41.19$, $p < .0001$, showing interference from the irrelevant target level, as indexed by faster responses to congruent trials compared to the responses on incongruent trials (594 ms vs. 646 ms). As anticipated, the significant Age Group x Congruency interaction revealed a decrease of interference with advancing age (100 ms in 7-year-olds vs. 36 ms in 11-year-olds vs. 16 ms in 21-year-olds), $F(2, 58) = 9.10$, $p < .0001$. Subsequent contrast analyses revealed that adult level of performance was attained in 11-year-olds.

The main effect of Level Type reflected that shifting between target levels was associated with a cost, as indicated by faster responses on repetition trials compared to alternation trials (564 ms vs. 675 ms), $F(1, 58) = 152.55$, $p < .0001$. As expected, Age

Groups differed with respect to this effect, as indicated by a decrease of level-shifting costs with age (196 ms in 7-year-olds, 79 ms in 11-year-olds, and 57 ms in 21-year-olds), F (2, 58) = 22.42, p < .0001[1]. Subsequent contrast analyses revealed that adult level of performance was attained in 11-year-olds. In addition, interference on local-directed trials was more pronounced (73 ms) than on global-directed trials (31 ms; Target Level x Congruency; F (1, 58) = 9.67, p = .003). Finally, level-shifting costs on incongruent trials were higher (135 ms) than on congruent trials (86 ms; Level Type x Congruency, F (1, 58) = 9.63, p = .003). Importantly, the Age Group x Level Type interaction was qualified by a higher-order interaction including Congruency, F (2, 58) = 3.51, p = .036. In all age groups a higher cost was associated with level-shifting on incongruent trials than with level shifting on congruent trials (see Figure 3). Follow-up contrasts indicated that the increase in level-shifting costs on incongruent trials discriminated significantly between the child groups (p < .001), but not between the older children and adult participants.

5.3   Experiment II

The main goal of Experiment II was to examine developmental change in global precedence by using symbolic cues to indicate the level of analysis (global vs. local) in responding to the target stimulus. In addition, the results of Experiment II allowed us to assess the robustness of the developmental trajectories that emerged in the previous experiment. The symbolic cue used in Experiment II to indicate the "respond-to-global" task was a cartoon of an elephant and a cartoon of a mouse was used to indicate the "respond-to-local" task. The size of the cues was 4 cm x 2.5 cm. See Figure 4 for a schematic of the stimulus and cue display. All other details were identical to Experiment I.

5.3.1   Method
5.3.1.1   *Participants*

Participants in three age groups (7-year-olds, 11-year-olds, 21-year-olds) participated in the study. Children were recruited by contacting regular schools; the 21-year-olds were university students. Informed consent was obtained. The children received a small present for their participation and the 21-year-olds received course credits. All participants reported to be healthy, and had normal, or corrected-to-normal, vision. In order to obtain an estimate of the participants' intelligence scores, the child groups took the Raven SPM

---

[1] The finding that adult responses to alternation trials were both more accurate and slower may suggest a trade-off between speed and accuracy. Therefore, for this age group, we performed a follow-up analysis. We split (median) the participants into a low error group (N = 10) and a high error group (N = 11). Subsequently, a 2 (Error Group) x 2 (Trial Type) x 2 (Target Level) x 2 (Congruency) resulted in nonsignificant interactions with Error Group (Note: F < 1 in the Error Group x Trial Type interaction). Therefore, we discarded the apparent speed-accuracy trade-off in the Trial Type effects.
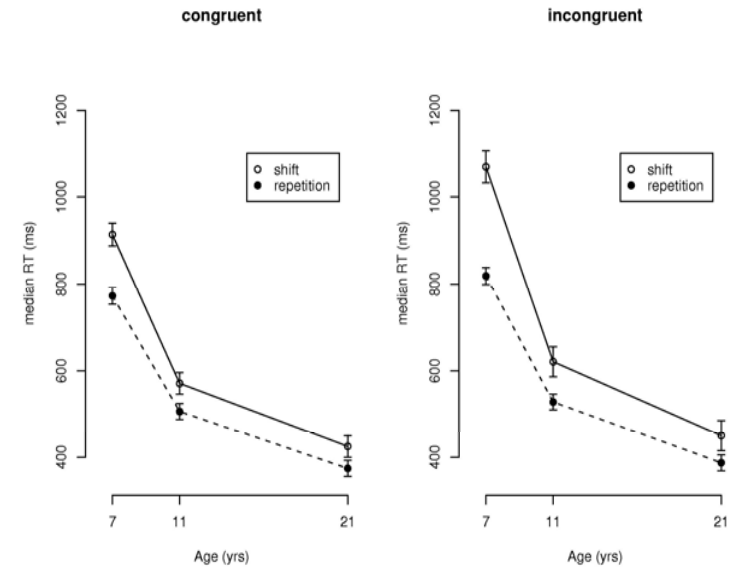
Figure 3. Median RTs (ms) as a function of Level Type for each age group, on congruent and on incongruent trials of Experiment I. "Shift" refers to shifting choice tasks (i.e., global vs. local or vice versa); "Repetition" refers to repeating choice tasks

and the 21-year-olds performed the Raven Advanced Progressive Matrices (APM; Raven, Raven, & Court, 1998). Scores were converted to quartile scores, following the norms for each age group. There was no significant difference between age groups on Raven quartile, F (2, 57) = 1.18, p = .316. Chi-square analyses indicated that gender distribution did not differ significantly between age groups, χ^2 (2) = .48, p = .786. The distribution of gender, age, and Raven scores across age groups are reported in Table 1.

5.3.2   Results

As above, the results are presented in two sections. The first section focuses on the performance on the single-level blocks, and the second section focuses on performance in mixed-level blocks. In the analyses reported below, Age Group (7-year-olds, 11-year-olds, 21-year-olds) served as between-subjects factor. The first five trials in each task block, trials with RTs shorter than 120 ms were excluded from the analyses. In addition,
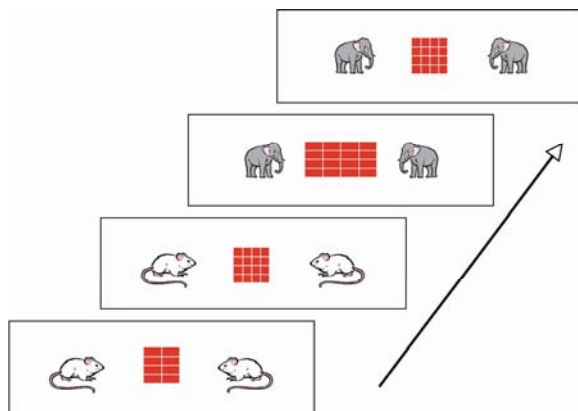
Figure 4. The figure displays a trial sequence proceeding from bottom to top, in a mixed-level block. The mice flanking the target served as cues that instructed the participant to perform the local task; the elephants flanking the target served as cues that instructed the participant to perform the global task. Before the task block started, participants were instructed to press the left key when the target consisted of small squares and to press the right key when it consisted of small rectangles, when the target was flanked by mice; and to press the left key when the target was a small square and to press the right key when it was a big rectangles, when the target was flanked by elephants. After each response, a response-cue interval occurred, followed by the presentation of a cue, and after another interval the target appeared. The first display is the last event of the previous trial, which required local discrimination. The second display indicates a task repetition, and the third display indicates a task shift to global discrimination

error trials and trials following an error were excluded from the RT analyses.

5.3.2.1   Single-level blocks

The square root of error percentages and median RTs were submitted to separate repeated measures ANOVAs, with Target Level (global vs. local) and Congruency (congruent vs. incongruent) as within-subjects factor. Twelve 7-year-olds were excluded from the analyses due to excessive error rates in one of the conditions. That is, these children performed extremely poorly (i.e., > 80% errors) on incongruent trials (six children failed on incongruent global-oriented trials and six children failed at incongruent local-oriented

trials). The excluded children did not differ from the 7-year olds, who remained in the analyses, in terms of gender or Raven score.

Errors. The ANOVA revealed a significant main effect of Age Group, showing a decrease in erroneous responding with advancing age (3.9% in 7-year-olds, 3.9% 11-year-olds, and 1.6% 21-year-olds), $F(2, 51) = 3.36$, $p = .043$. In addition, there was a significant main effect of Congruency, which consisted of fewer errors on congruent trials compared to incongruent trials (1.4% vs. 5.2%), $F(1, 51) = 38.35$, $p < .0001$. There was no difference between Age Groups with respect to this effect. All other effects, including the Target Level effect, fell short of significance.

Response latencies. The ANOVA yielded a main effect of Age Group; the speed of responding increased with advancing age (810 ms in 7-year-olds, 559 ms in 11-year-olds, and 401 ms in 21-year-olds), $F(2, 51) = 80.14$, $p < .0001$. Moreover, a significant effect of Target Level, reflecting global precedence, involved faster responses to global-directed trials compared the responses to local-directed trials (567 vs. 613 ms), $F(1, 51) = 10.31$, $p = .002$. As in Experiment I, there was no difference between Age Groups with respect to this effect ($F < 1$). Finally, we found a main effect of Congruency indicating interference of the non-target level, as shown by faster responses to congruent trials compared to the responses to incongruent trials (569 vs. 611 ms), $F(1, 51) = 26.52$, $p < .0001$, and a trend to an interaction with Age Group was observed, $F(2, 51) = 2.56$, $p = .088$. All other effects were not significant.[2]

5.3.2.2   Mixed-level block

The square root of the error percentages and median RTs were submitted to separate repeated measures ANOVAs with Target Level (global vs. local), Congruency (congruent vs. incongruent), and Level Type (level repetition vs. level alternation) as within-subjects factors. Four additional participants (three 7-year-olds and one 11-year-old) were excluded from the analysis due to extremely high error rates (i.e., > 80%) on incongruent trials (two 7-years-olds failed on incongruent local-directed trials, one 7-year-old and one 11-year-old failed on incongruent global-directed trials). Thus, data of the remaining nine 7-year-olds, twenty 11-year-olds, and twenty-two 21-year-olds were included in the analyses.

---

[2] One might argue that youngest age group did not include sufficient participants (i.e., 12) to perform a reliable analysis of Age Group effects. The pattern of findings did not change, however, when the data from the youngest age group were excluded from the analysis. That is, for errors, we obtained significant main effects for Age Group ($F(1, 40) = 7.27$, $p = .010$), and Congruency ($F(1, 40) = 58.10$, $p < .0001$), all other effects failed to reach significance. For RT, we observed significant main effects for Age Group ($F(1, 40) = 79.18$, $p < .0001$), Target Level ($F(1, 40) = 33.35$, $p < .0001$), and Congruency ($F(1, 40) = 44.07$, $p < .0001$), all other effects failed to reach significance.

Figure 5. Median RTs (ms) as a function of Level Type for each age group, on congruent and on incongruent trials of Experiment II. "Shift" refers to shifting choice tasks (i.e., global vs. local or vice versa); "Repetition" refers to repeating choice tasks. NB. N=9 in 7-year-olds, see text for further clarification. The interaction effect was significant in the two oldest age groups, for an indication of the performance of 7-year-olds, developmental course in gray

congruent and incongruent trials was 2% in Experiment I vs. 4.5% in Experiment 2).

Response latencies. The ANOVA of median RTs revealed a significant main effect of Experiment (F (1, 80) = 15.42, p < .0001; RT = 436 ms in Experiment I vs. 480 ms in Experiment II). Moreover, the interaction of Experiment and Target Level was significant. The global-precedence effect was 78 ms in Experiment I vs. 46 ms in Experiment II, F (1, 80) = 9.35, p = .003.

5.4.2 Mixed-level blocks

The square root of errors and median RTs were submitted to a repeated measures ANOVA, with Target Level (global vs. local), and Congruency (congruent vs.

incongruent), and Level Type (level repetition vs. level alternation) as within-subjects factors.

Errors. The ANOVA of errors revealed a significant main effect of Experiment (F (1, 79) = 33.44, p < .0001; 2% in Experiment I vs. 6% in Experiment II). In addition, there was a significant interaction of Experiment, Congruency, and Level Type, F (1, 79) = 12.19, p = .001. This interaction revealed elevated level of errors on incongruent trials in both Experiments, however the difference between repetition and alternation trials was only significant on incongruent trials in Experiment II (8%).

In addition, we found a significant interaction of Age Group, Experiment, and Level Type that revealed an elevated level of errors on level-alternation trials in 21-year-olds in Experiment I. However the remaining differences between level-repetition and level-alternation trials in both age groups were non-significant in both experiments, F (1, 79) = 4.24, p = .043.

Response latencies. The ANOVA of median RTs revealed a significant main effect of Experiment (F (1, 79) = 31.82, p < .0001; RT = 483 ms in Experiment I vs. 596 ms in Experiment II). Further, we observed a higher-order interaction between Age Group, Experiment, and Congruency (F (1, 79) = 4.83, p = .031). Follow-up analyses revealed that the interference effect was considerably larger in Experiment 2 compared to Experiment 1, in particular for children (Experiment I: 36 ms in 11- year-olds vs. 19 ms in 21-year-olds; Experiment II: 116 ms in 11- year-olds vs. 49 ms in 21-year-olds).

A three-way interaction of Age Group, Experiment, and Level Type approached significance, F (1, 79) = 3.84, p = .053. Subsequent analyses revealed that in both Age Groups and Experiments shifting between levels was associated with a cost in RT, as reflected by larger RTs on alternation trials compared to repetition trials. These level-shifting costs did not differ significantly between age groups in Experiment I; in Experiment II shift costs they were, however, significantly larger in 11- year-olds (244 ms) than in 21-year-olds (149 ms).

Finally, the interaction of Experiment, Congruency, and Level Type was significant, F (1, 79) = 7.06, p = .010. Level-shifting costs were more pronounced on incongruent trials compared to congruent trials, and this difference was larger for children compared to adults (11-year-olds: 120 ms on congruent trials vs. 195 ms on incongruent trials; 21-year-olds: 92 ms on congruent trials vs. 115 ms on incongruent trials) and for Experiment I compared to Experiment II (Experiment I: 58 ms on congruent trials vs. 77 ms on incongruent trials; Experiment II: 156 ms on congruent trials vs. 237 ms on incongruent trials).

In conclusion, the data obtained in Experiments I and II are consistent. In general, the global precedence effect was present in both experiments, and was more pronounced in Experiment I. In addition, the effect of Congruency was present in both experiments, and was however more pronounced in Experiment II.

5.5   Discussion

The current study set out to address the developmental pattern of the ability to shift attention between global and local levels of hierarchical stimuli. Participants in three age groups (7-year-olds, 11-year-olds, and 21-year-olds) performed on a task that allowed for the examination of the development in 1) the direction of attention to either global or local levels of a stimulus (i.e., global or local preference); 2) the susceptibility to interference; and 3) the flexible direction of attention to global or local levels of hierarchical stimuli. We applied the global-local paradigm, including target stimuli adapted from Kimchi (1992; Kimchi & Palmer, 1985), and adopted a hybrid version of the "alternating-runs" and the "explicit task-cuing" procedures from the task-switching literature (e.g., Meiran, 1996; Rogers & Monsell, 1995). The present study involved two experiments. In the first experiment participants were presented with task cues that allowed for an easy match to the target stimulus, whereas in the second experiment target level was indicated by a symbolic cue.

We consider first the adult findings, which should provide the context for interpreting the pattern of developmental findings. Given the recurrent observation of a global precedence effect in adult studies that involved identical target stimuli to the current study (e.g., Kimchi, 1992; Kimchi & Palmer, 1985), we expected a global precedence effect. In both experiments, we observed a global precedence effect. Responses to global-oriented trials were faster and more accurate compared to the responses to local-oriented trials. This difference was observed in both single and mixed trial blocks. This finding is consistent with a host of previous studies showing that the perception of global level of a hierarchical stimulus is primary (e.g., Enns et al., 2000; Goto et al., 2004; Miller & Navon, 2002; Navon, 2003). In addition, we expected interference of irrelevant information to occur on trials that required a different response to the relevant and the irrelevant target level (i.e., incongruent trials). Again, in both experiments and on both the single- and mixed-level block, responses on incongruent trials were slower and less accurate compared to responses congruent trials. In addition, this interference effect was asymmetrical. That is, responses to the local target level were more affected by irrelevant information from the non-target level than were responses to the global target level of the hierarchical stimuli (see Kimchi, 1992; Navon, 1977). Finally, when participants were required to shift between target levels, we expected a level-shifting effect to occur (i.e., we expected responses to trials that involved level repetition to be faster and more accurate than responses to trials that involved level alternation). Indeed, in both experiments, level-shifting responses were slower and less accurate than level-repeating responses. This finding is consistent with results reported previously in the global-local and task-switching literatures (e.g., Hübner, 2000; Monsell,

2003; Shedden et al., 2003). Finally, the costs of level-shifting were larger on incongruent trials than they were on congruent trials. That is, responses are slower when the level at which the target is presented on the current trial is different from the previous level, and this cost is larger when the target and the information presented at the to-be-ignored level relate to competing responses. This pattern of results is consistent with previous findings showing that the size of level-repetition effects depends on competition from the ignored level (Shedden et al., 2003).

We now consider developmental change in the processing of hierarchical stimuli. Our hypotheses of developmental change were guided by the "orthogenetic principle" (Werner, 1948, 1957), which stipulates that processing of hierarchical stimuli in children follows a global-to-local pathway as the children grow older. That is, young children process a hierarchical stimulus by directing their attention initially in an undifferentiated state (i.e., the global whole), but with increasing age children develop towards a coordinated integration of specialized components (i.e., the local parts; Enns et al., 2000). One goal of the present study was to examine the development of global-local processing, when children were required to direct their attention either to the global or the local level of a hierarchical stimulus. We anticipated that global precedence would occur, which is strongest for the youngest children, and decreases with advancing age. The current results are largely consistent with our prediction. In both experiments, all age groups showed a global precedence effect, and, importantly, the strength of the global precedence effect decreased with age, albeit on mixed-level blocks only. A similar pattern of results was obtained for stimulus congruency. Based on findings reported in the ANT literature (Konrad et al., 2005; Rueda et al., 2004) and related literatures (e.g., Bunge et al., 2002; Ridderinkhof & Van der Molen, 1995), it was expected that the susceptibility to interference of the irrelevant stimulus level would decrease with advancing age. Indeed, the interference effect decreased in strength as children grow older, but, similar to the global precedence effect, this age-related trend was observed only for the performance on mixed-level blocks, not on single level blocks. Both effects are considerably stronger for mixed-level blocks compared to single-level blocks (respectively, 101 ms vs. 76 ms for the global precedence effect in Exp. I and 76 vs. 46 ms in Exp. II; 52 ms vs. 27 ms for the interference effect in Experiment I and 86 ms vs. 42 ms in Experiment II). Thus, the most parsimonious interpretation is that on mixed-level blocks there was simply more room for developmental trends to occur.

Finally, the global-local paradigm allows for the examination of developmental change in the ability to flexibly shift attention between target levels. In both experiments, we observed that the costs related to level-shifting decreased with increasing age, reaching the adult level of performance at age 12. This finding is consistent with results reported previously in the developmental literature on task-switching in children (e.g., Cepeda et al., 2001; Crone et al., 2006; Huizinga et al., 2006). The age-related decrease

in the costs involved in level-shifting was not altered by target level. This finding is consistent with previous findings indicating that the costs associated with shifting from local-to-global and global-to-local are symmetrical (e.g., see also Lamb & Yund, 1996; Robertson, 1996; Ward, 1982; but see Shedden et al., 2003). Congruency, in contrast to Target Level, did alter the age-related change in level-shifting. That is, the youngest children in Experiment I suffered more from level-shifting on incongruent than on incongruent trials. This finding supports the notion that young children are more susceptible to interference from the to-be-ignored level (e.g., Mondloch et al., 2003).

Before closing, we have to address the results that emerged from the consistency check performed on the data from the two experiments. The experiments were identical except for the cues used to indicate target level. In Experiment I, the cues were shapes (squares or rectangles) of a size corresponding to that of the respective target pattern. It was assumed that these cues allowed participants to focus attention to the specified size in advance, and thus would make the global-local task doable even for young children. In Experiment II, the cues were cartoons of an elephant (indicating the global level) and of a mouse (indicating the local level). In this experiment, participants had to translate the cartoon into a category code indicating target level. In addition, the cue in Experiment 1 repeated the target-response mapping whereas the cue in Experiment II did not contain this information. Overall, the outcomes of the consistency check indicated that the basic pattern of findings was very similar in both experiments. Both experiments produced a global precedence effect (somewhat smaller in the first than in the second experiment), an interference effect (considerably stronger for the second compared to the first experiment), and a level-repetition effect. The major difference between experiments referred to the performance of the youngest child group (i.e., the 7-years olds). These children were able to perform well in Experiment I, but a considerable number of their peers (about half of the group) were not able to perform their task. More specifically, the children, who failed this task, committed unacceptable numbers of errors on incongruent trials. We don't have a ready explanation for their failure. The children, who failed, did not differ from their peers in terms of age or intellectual level as indexed by their Raven scores. Moreover, the task demands on single-level blocks did not differ between experiments, as the cue is not needed for directing attention. Yet, the 7-years olds failed on both single- and mixed-level blocks. Finally, their performance was reasonable on congruent trials but went astray on incongruent trials. This difference suggests that they are able to process the symbolic cues, but not under all conditions. A tentative interpretation of the trial-specific failure of the youngest child group refers to an overload of working memory. In Experiment II, participants need to retain the target-response mapping in working memory, whereas in Experiment I the mapping is refreshed on each trial. In Experiment II, participants needed to translate the cartoon into a category code, whereas in Experiment I they could use the size of the cue for directing attention to the

corresponding level. Although participants could ignore the cue on single-level trials, it cannot be excluded that they, nevertheless, continued to pay attention to the cue. In this regard, it seems fair to say that the experiments differ in terms of the demands on working memory. Possibly, the demands on working memory exceeded the capacity in the youngest children. The selective failure on incongruent trials can then be explained by refering to notions suggesting that the ability to inhibit is inversely proportional to the availability of working memory (e.g., Bjorklund & Harnishfeger, 1990; Kipp, 2005). On incongruent trials, the target at the to-be-attended level and the information at the to-be-ignored level activate competing responses (e.g., Ridderinkhof & Van der Molen, 1995) and participants must inhibit the imminent response to the information at the to-be-ignored level. Inhibition is more difficult for the youngest children (Van den Wildenberg & Van der Molen, 2004), in particular when their working memory is loaded (e.g., Eenshuistra, Ridderinkhof, & Van der Molen, 2004).

In conclusion, the current study aimed to examine the development of 1) the direction of attention to either global or local levels of a stimulus (i.e., global or local preference); 2) the susceptibility to interference; and 3) the flexible direction of attention to global or local levels of hierarchical stimuli. The results indicated a global precedence effect in all age groups, and a trend towards local processing. In addition, the susceptibility to irrelevant information decreased with age, and the ability to flexibly shift between global-local levels increased. Both abilities attained young-adult level in 11-year-olds. Thus, given the robust and consistent findings obtained in the current study, the global-local paradigm seems to provide a complementary procedure to the strategy adopted in the recent ANT studies (e.g., Fan et al., 2002; Konrad et al., 2005; Rueda et al., 2004). In future research, it would be of interest to employ neuro-imaging to evaluate the biological plausibility of the attentional functions of interest in the current study, and their development in terms of functionally and anatomically distinct systems.

Chapter 6 ● Age-related change in executive function: Developmental trends and a latent variable analysis

Abstract

This study examined the developmental trajectories of three frequently postulated executive function (EF) components, Working Memory, Shifting, and Inhibition of responses, and their relation to performance on standard, but complex, neuropsychological EF tasks, the Wisconsin Card Sorting Task (WCST), and the Tower of London (ToL). Participants in four age groups (7-, 11-, 15-, and 21-year-olds) carried out nine basic experimental tasks (three tasks for each EF), the WCST, and the ToL. Analyses were done in two steps: (1) analyses of (co)variance to examine developmental trends in individual EF tasks while correcting for basic processing speed, (2) confirmatory factor analysis to extract latent variables from the nine basic EF tasks, and to explain variance in the performance on WCST and ToL, using these latent variables. Analyses of (co)variance revealed a continuation of EF development into adolescence. Confirmatory factor analysis yielded two common factors: Working Memory and Shifting. However, the variables assumed to tap Inhibition proved unrelated. At a latent level, again correcting for basic processing speed, the development of Shifting was seen to continue into adolescence, while Working Memory continued to develop into young-adulthood. Regression analyses revealed that Working Memory contributed most strongly to WCST performance in all age groups. These results suggest that EF component processes develop at different rates, and that it is important to recognize both the unity and diversity of EF component processes in studying the development of EF.

6.1 Introduction

Across development, children become increasingly more able to control their thoughts and actions (for a review see: Diamond, 2002). This change has been associated with the development of executive function (EF), which is an umbrella term for various cognitive processes that subserve goal-directed behavior (Miller & Cohen, 2001; see also Luria, 1966; Shallice, 1982). EF is especially important in novel or demanding situations (Stuss, 1992), which require a rapid and flexible adjustment of behavior to the changing demands of the environment (Zelazo, Muller, Frye, & Marcovitch, 2003). EF is thought to rely strongly on prefrontal cortex (PFC), as indicated by studies showing that patients with lesions to PFC perform poorly on tasks such as the Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948) and the Tower of London (ToL; Shallice, 1982; for a

review see: Stuss & Levine, 2002). PFC patients typically perseverate, i.e., they persist in sorting according to the rule that was previously correct (e.g., Anderson, Damasio, Jones, & Tranel, 1991; Milner, 1963; Nagahama, Okina, Suzuki, Nabatame, & Matsuda, 2005; Stuss et al., 2000). On the ToL, which requires spatial problem solving by moving balls in order to reach a pre-specified goal, PFC patients require more moves to solve the problem (e.g., Andres & Van der Linden, 2001; Carlin et al., 2000; Morris, Ahmed, Syed, & Toone, 1993; Owen, Downes, Sahakian, Polkey, & Robbins, 1990).

Children show a similar pattern as patients with PFC damage; that is, they also perseverate on the WCST and require more moves to solve ToL problems (Anderson, Anderson, & Lajoie, 1996; Baker, Segalowitz, & Ferlisi, 2001; Chelune & Baer, 1986; Chelune & Thompson, 1987; Heaton, Chelune, Talley, Kay, & Curtis, 1993; Kirk & Kelly, 1986; Lehto, 2004; Lehto, Juujaervi, Kooistra, & Pulkkinen, 2003; Paniak, Miller, Murphy, & Patterson, 1996; Welsh, Pennington, & Groisser, 1991). The slow development of EF has been attributed to the protracted maturation of PFC (e.g., Diamond, 2002). Conclusive evidence about the developmental trajectories of the different EF components in relation to the performance on standard neuropsychological EF tasks has yet to be established. In this study, we examined the development of EF component processes by using a multi-group confirmatory factor analysis. Where we have at our disposal multiple indicators of a given latent variable, this approach has the advantage that it allows us to study performance at the level of the latent variables, according to a pre-specified model of EF.

6.1.1 Decomposition of executive function

A major theoretical issue concerns the organization of EF. It has been suggested that EF is unitary, i.e., that it does not include distinct sub-functions or sub-components. This means that the cognitive and behavioral impairments seen after PFC damage can be explained entirely in terms of one dysfunctional system (e.g., Cohen & Servan-Schreiber, 1992; Duncan, Emslie, Williams, Johnson, & Freer, 1996; Kimberg, D'Esposito, & Farah, 1997). For example, Kimberg et al. (1997) posited that all deficits in PFC function can be attributed to deficits in working memory. In contrast, others view EF as multi-faceted (non-unitary). These authors argued that EF involves several discrete cognitive processes that have a relatively focal neural representation (e.g., Baddeley, 1986; Stuss, Shallice, Alexander, & Picton, 1995; see also Teuber, 1972). The multi-faceted nature of EF is suggested by behavioral studies incorporating batteries of widely used EF tasks. These studies yielded low or nonsignificant correlations between tasks and exploratory factor analysis yielded multiple factors (Brocki & Bohlin, 2004; Culbertson & Zillmer, 1998; Lehto, 1996; Levin et al., 1996; Pennington, 1997; Robbins et al., 1994; Welsh et al., 1991).

Neuroimaging studies provide evidence in support of the multi-faceted nature of EF, as different components of EF are seen to rely on different parts of PFC. For example, the ability to maintain information in working memory has been found to recruit mostly lateral PFC (Narayanan et al., 2005; Smith & Jonides, 1999). In contrast, switching between tasks is thought to rely on medial PFC (Crone, Wendelken, Donohue, & Bunge, 2006; Rushworth, Walton, Kennerley, & Bannerman, 2004). Finally, the ability to inhibit responses was found to rely on orbitofrontal cortex (e.g., Aron, Robbins, & Poldrack, 2004; Roberts & Wallis, 2000). Thus, different regions within PFC subserve different components of goal-directed behavior.

At this point, it should be noted that the problem of "task impurity" hinders the interpretation of results reported in behavioral and neuroimaging studies using multiple EF tasks. Task impurity refers to the fact that a single indicator (operationalization) of a given construct (e.g., Working Memory) can rarely, if ever, be viewed as a pure measure of that construct. Most measures are contaminated by random error and systematic error (see Kline, 1998, p.55). The task impurity problem is highly relevant to EF research, as the manifestation of EF components invariably involves other (non-EF) cognitive processes (e.g., Miyake et al., 2000).

Miyake and colleagues (2000) presented one way to address the task impurity problem. They proposed using multiple tasks to measure each EF component and adopting a latent variables approach to extract the variance common to those tasks. Latent variables (as incorporated in structural equation models; SEM) refer to what is shared among tasks that are assumed to tap a given EF. The latent variable approach minimizes the task impurity problem, and is therefore especially informative in developmental studies (e.g., Nunally & Bernstein, 1994, p.85). Using confirmatory factor analysis, Miyake et al. (2000) examined the separability of three frequently postulated EF components: "Working Memory", "Shifting", and "Response Inhibition" (hence forth: Inhibition). Miyake et al. (2000) focused on these three EF components because: 1) they are well-circumscribed, lower-level functions that can be operationalized in a fairly precise manner; 2) they can be studied using commonly used tasks; and 3) they have been implicated in the performance of more complex EF tasks, such as the WCST and ToL. Miyake et al. (2000) tested healthy young-adults on multiple tasks tapping Working Memory, Shifting, and Inhibition, and several standard but complex neuropsychological tasks, including the WCST and the Tower of Hanoi (similar to the ToL). The results showed that, although moderately correlated, Working Memory, Shifting and Inhibition were separable constructs (see also Fisk & Sharp, 2004). Moreover, the EF component processes differentially predicted performance on the complex neuropsychological tasks. For example, Shifting predicted WCST performance, whereas Inhibition predicted ToH performance.

6.1.2   Development of executive function

Developmental studies using standard neuropsychological tasks have shown that EF has a protracted course of development, beginning in early childhood and continuing into adolescence. However, these EF tasks are subject to distinct developmental trajectories. For example, on the WCST, analysis of perseverative errors indicates that the performance of children is comparable to that of young adults by 12 years of age; however, analysis of failure-to-maintain set indicates that children do not reach adult levels of performance until 13-15 years of age (e.g., Chelune & Baer, 1986; Chelune & Thompson, 1987; Levin et al., 1991; Welsh et al., 1991). Similarly, on the ToL task, performance based on errors appears to continuously improve from middle childhood into young adulthood; however, when performance is based on both errors and time, adult levels of performance may be attained as early as 13 years of age (Baker et al., 2001; see also Levin et al., 1996).

There is a growing body of research indicating differential trends in the development of EF component processes[1]. These studies, although not entirely unequivocal, show that adult-level performance on different EF tasks is attained at different ages during childhood and adolescence (for reviews see: Diamond, 2002; Welsh, 2002). Working memory capacity has been found to gradually develop throughout childhood and into adolescence (e.g., Beveridge, Jarrold, & Pettit, 2002; Brocki & Bohlin, 2004; DeLuca et al., 2003; Gathercole, Pickering, Ambridge, & Wearing, 2004; Hitch, Halliday, Dodd, & Littler, 1989; Luciana, Conklin, Hooper, & Yarger, 2005; Luciana & Nelson, 1998; Luna, Garver, Urban, Lazar, & Sweeney, 2004). In addition, recent studies on the development of task shifting abilities all show that the cost of switching between tasks decreases as children grow older, with adult levels of performance being attained around the age of 12 (Cepeda, Kramer, & Gonzalez de Sather, 2001; Crone, Bunge, Van der Molen, & Ridderinkhof, 2006; Huizinga & Van der Molen, 2005a; Kray, Eber, & Lindenberger, 2004). Finally, inhibitory control was found to increase throughout childhood (e.g., Klenberg, Korkman, & Lahti Nuuttila, 2001), and to reach adult-level of performance in late childhood, around the age of 12 (Bédard et al., 2002; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Durston et al., 2002; Ridderinkhof & Van der Molen, 1995; Van den Wildenberg & Van der Molen, 2004), or early adolescence (Williams, Ponesse, Schachar, Logan, & Tannock, 1999).

---

[1] A growing body of research appeared recently, focusing on EF in pre-school aged children, (e.g., Carlson, Mandell, & Williams, 2004; Diamond, Briand, Fossella, & Gehlbach, 2004; Diamond, Kirkham, & Amso, 2002; Espy & Bull, 2005; Gerstadt, Hong, & Diamond, 1994; Kirkham, Cruess, & Diamond, 2003; Muller, Zelazo, Hood, Leone, & Rohrer, 2004). These studies, however, applied tasks suitable for very young children that cannot be compared to the tasks used for the age range examined in the current study without violating the assumption of measurement invariance (Meredith, 1993).

A straightforward interpretation of the developmental trends of the EF component processes is hampered by a number of factors. First, different tasks are used across studies to measure the same EF component. For example, the developmental trajectory of Inhibition has been assessed by using a Go/NoGo task (e.g., Durston et al., 2002), the Eriksen Flankers task (e.g., Bunge et al., 2002; Ridderinkhof & Van der Molen, 1995), and the Stop-signal task (e.g., Bédard et al., 2002; Van den Wildenberg & Van der Molen, 2004; Williams et al., 1999). Second, it is unclear whether children at various ages use the same strategy when performing on EF tasks. This issue concerns the question of measurement invariance, i.e., whether we are actually measuring the same construct across age (Meredith, 1993). Third, many developmental studies focused on a single EF component process. This precludes a reliable assessment of developmental change across EF component processes, because differential rates might be due to different samples rather than components. Thus, a reliable assessment of the developmental patterning of EF component processes requires homogeneous age groups and the application of a latent variables approach to extract what the various tasks used to tap EF component processes have in common. This approach has been adopted by Lehto et al. (2003), who were first to assess the patterning of EF component processes in children by using SEM. Importantly, they observed the same factor structure in a group of 8- to 13-year-old children as previously found by Miyake et al. (2000) in adults. In the present study, we hope to contribute to these results by adopting a multi-group design, and thus providing a more graded assessment of developmental change in EF component processes.

6.1.3   The present study

In the present study, we adopted the conceptual framework of Miyake et al. (2000) to assess developmental change in EF. The main goal of this study was to examine age-related changes in the three EF components distinguished by Miyake et al. (2000), i.e., Working Memory, Shifting, and Inhibition. In order to shed some light on the development of these EF component processes, we tested children in three homogeneous age groups (i.e., 7-year-olds, 11-year-olds, 15-year-olds), in addition to a group of young-adults (i.e., 21-year-olds). Adult level on EF tasks is typically reached in late childhood or early adolescence (for reviews see: Diamond, 2002; Welsh, 2002). The decision to limit the youngest group to 7 years olds was based on the consideration that the present task battery was probably too difficult for children younger than 7 years of age. The EF components, Working Memory, Shifting, and Inhibition, were indexed by nine experimental tasks, three for each EF component.

Working Memory was defined as the collection of cognitive processes that temporarily retain information in an accessible state, suitable for carrying out any mental

task (Cowan, 1998). The essence of this component is the monitoring and coding of incoming information with respect to relevance and the replacement of information that is no longer relevant by newly relevant information. Shifting was interpreted as shifting back and forth between multiple tasks (Allport, Styles, & Hsieh, 1994; Monsell, 1996, 2003). When different (usually choice RT) tasks are mixed within blocks, shifting between tasks typically results in an increase in RT and a decrease in accuracy (i.e., shift costs). Inhibition was conceptualized as the ability to deliberately inhibit dominant, automatic, or pre-potent responses (Logan & Cowan, 1984). The essence of this EF component lies in the suppression of a response or in the control of interfering stimuli or competing responses.

We adopted two approaches in analyzing the data. First, we conducted a standard analysis of variance approach. Second, we took a latent variable approach, i.e., multi-group confirmatory factor analyses (Dolan, 2000; Meredith, 1993). We examined i) the organization of executive function in children and young-adults by investigating whether the indicators of the Working Memory, Shifting, and Inhibition tasks measured the same constructs across age, ii) whether this organization changes across development, and iii) how EF component processes contribute to the performance on the WCST and the ToL across age groups, again following Miyake et al. (2000). We included the WCST and the ToL because these tasks have been used previously to study the development of EF (Anderson, Byrd, & Berg, 2005; Baker et al., 2001; Chelune & Baer, 1986; Lehto, 2004; Welsh et al., 1991).

6.2   Method
6.2.1   Sample

The present study included four age groups: 71 7-year-olds (39 female, M age = 7.2 (age range: 6-8); M Raven-quartile = 3.6 (SD = 0.88); M number of years of education = .56 (SD = .13)), 108 11-year-olds (62 female, M age = 11.2 (age range: 10-12); M Raven-quartile = 3.2 (SD = 0.93); M number of years of education = 3.92 (SD = .13)); 111 15-year-olds (58 female, M age: 15.3 (age range: 14-16); M Raven-quartile = 3.1 (SD = 0.99); M number of years of education: 7.20 (SD = .20)), and 94 21-year-olds (72 female, M age = 20.8 (age range 18-26); M Raven-quartile: 3.7 (SD = 0.52); M number of years of education = 10.55 (SD = .31)). Children were recruited from regular local schools; the 21-year-olds were university students. Teachers assisted in the selection process in order to exclude children with health problems, neurological damage, or psychiatric problems. Similar information was derived from a self-report of the 21-year-olds. Informed consent was obtained. All participants had normal, or corrected-to-normal, vision. The 7- and 11-year-olds received a small present for their participation, the 15-year-olds received € 10, and the 21-year-olds received course credit.

To assess intelligence, we used a non-verbal IQ test, the Standard Progressive Matrices (SPM; Raven, Court, & Raven, 1985). Raw scores were converted to quartile scores, following the norms for each age group. Mean Raven-quartile scores differed between groups F (3, 359) = 11.77, p < .001. Post-hoc Bonferroni tests indicated significant differences between the young-adult group and the other groups. Preliminary analyses were performed to control for IQ, and we found the relationship between IQ and the different EF measures absent both within-groups and between-groups. This is consistent with other studies that found no relation between IQ and EF (e.g., Anderson et al., 1991; Bechara & Martin, 2004; but see Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Conway, Kane, & Engle, 2003; Duncan et al., 1996 for opposite results). The gender distribution across groups differed significantly, χ^2 (3) = 14.64, p = .002. This was caused by a relatively larger proportion of women in the young-adult group. The addition of gender as a covariate in the analyses reported below did not change any of the main effects or interactions involving the respective task manipulations. Therefore, Raven IQ scores and gender were not included in the analyses reported below.

6.2.2 Experimental tasks

The experimental tasks were designed to assess the EF components Working Memory, Shifting, and Inhibition. The task battery also included three complex EF tasks WCST, the ToL, and the Random Number Generation task. The results relating to the last task are presented elsewhere (Huizinga & Van der Molen, 2005b). Task administration was computerized, and presented on a Toshiba Satellite 1600 laptop (Intel Celeron 800 mHz processor; 15 inch 60 Hz monitor). All tasks required only left- and right-hand responses. The response button for the right hand was the "?" key on the computer keyboard, for the left hand it was the "z" key (responses were counterbalanced across participants).

The nine EF tasks that we used were speeded choice reaction time (RT) tasks. The WCST and ToL were computerized versions of the standard neuropsychological tests. The participants were coached to balance speed and accuracy when responding on all tasks, except the WCST and ToL. Care was taken to ensure that participants understood the instructions, as indexed by verbal report, response accuracy, and stable RTs.

6.2.2.1 *Working memory*

Tic Tac Toe. In the Tic Tac Toe task (adapted from Milner, 1971), participants were required to retain visual information about the orientation of a pattern of figures active in their working memory. The task consisted of a memorizing phase and a recognition phase. A pattern consisting of *X*s and *O*s was presented within a 3x3 grid during the memorizing phase. Working memory load was varied by using patterns consisting of

---

three vs. four letters (i.e., low memory load vs. high memory load). The recognition phase was initiated by pressing the space bar. During this phase, the *X*s and *O*s were presented one after another at different positions in the grid, each for a period of 600 ms, in series that varied from four to seven presentations for the low memory load, and from four to nine presentations for the high memory load. As soon as the pattern of *X*s and *O*s included the pre-specified pattern indicated in the memorizing phase, participants were required to press a button. Participants first received a practice block of 3 trials, followed by one block including 15 trials for low memory load and one block including 15 trials for high memory load. The low and high memory load blocks were counterbalanced across participants to control for order effects. Participants had 3500 ms to respond; the time interval between trials varied randomly between 900 and 1100 ms (drawn from a uniform distribution). The main dependent variable was the proportion of correct trials.

Mental Counters. The Mental Counters task (adapted from Larson, Merritt, & Williams, 1988) required participants to retain numerical information active in their working memory. Participants had to keep track of the values of two or three (blocked) independent "counters", which change rapidly and in random order. The counters consisted of a horizontal line, above or below which squares appear. Participants were required to add 1 to the value of the counter, when a square appeared above the line, and to subtract 1, when it appeared below the line. When any counter reached a given criterion value, participants had to press a button. The series of stimuli presented comprised five or seven (chosen randomly and equiprobably); there were two blocks consisting of 15 response probes. Participants had 3500 ms to respond. The interval between consecutive presentations of squares varied randomly between 1000 and 1300 ms (drawn from a uniform distribution). The main dependent variable was the proportion of correct trials.

Running Memory (adapted from Kramer, Larish, & Strayer, 1995). The stimuli in this task were pictures of fruit (apple, banana, lemon, pair, grape) and of animals (cat, dog, bear, frog, mouse, rabbit, horse, monkey), taken from the Snodgrass & Vanderwart (1980) pictures set. The participants were presented with series of fruit-and-animal pairs, and every four to seven presentations (equally distributed within blocks) they were required to indicate by a button press whether or not the current pair matched the last presentation of the pair (e.g., apple - cat ?). The participants pressed one button on the computer keyboard if the pair matched the previous presentation of this pair (i.e., apple – cat) and another button if it did not match the last presentation (e.g., apple – frog). Feedback was presented following each response (i.e., '+' for a correct response, '-' for an incorrect response, 'x' for an omission), and this information remained on the screen for 400 ms. The association between particular pictures of fruit and animals changed over presentations. Working memory load was varied by increasing the number of animals in a run of trials (i.e., a low memory load vs. a medium memory load vs. a high memory load

was created by using, respectively, eight pictures of fruit and three of animals; vs. eight pictures of fruit and four of animals; vs. eight pictures of fruit and five of animals). The participants received one block for each condition, consisting of eight response probes. The order of the low, medium, and high memory load blocks was counterbalanced across participants to control for order effects. Participants had 3500 ms to respond. The time interval between the stimulus pairs varied randomly between 1000 and 1250 ms (drawn from a uniform distribution). The main dependent variable was the proportion of correct trials.

6.2.2.2   Shifting

Local-Global (adapted from Miyake et al., 2000). Participants responded to randomly presented rectangles or squares by pressing a left or right response button, respectively. Larger (global) rectangles/squares consist of smaller (local) rectangles or squares. In one block participants responded to the local figure, in a second block the responded to the global figure (blocks 1 and 2, in randomized order; 30 practice trials and 50 experimental trials per block), and in the third block they alternated between a series of four "local" trials and a series of four "global" trials (block 3; 90 practice trials, 150 experimental trials). A cue indicated to which dimension (global or local) the participants should respond. Cues that related to the global (local) dimension consisted of a big (small) square, presented at one side of the target stimulus, and a big (small) rectangle, presented at the other side of the target stimulus. The color of cues and target was red. They remained on the screen until a response was given. Participants had 3500 ms to respond. The time interval between presentation of the cue and of the target stimulus was 500 ms. The interval between the response and the presentation of the cue was fixed at 1000 ms. The main dependent variables were the median response latencies on task repetition and task alternation trials.

Dots-Triangles (derived from Rogers & Monsell, 1995). Varying numbers of either red dots or green triangles were presented in a 4x4 grid on the screen (i.e., three to eight dots or triangles per half of the grid; equally distributed). During the "dots" task, participants had to decide whether there are more dots in the left or the right part of the screen (block 1; 30 practice trials, 50 experimental trials). During the "triangles" task, participants had to decide whether there are more triangles in the top or in the bottom part of the screen (block 2; 30 practice trials, 50 experimental trials). Blocks 1 and 2 were administered in randomized order. In the third block (90 practice trials, 150 experimental trials), a series of four "dots" trials and a series of four "triangles" trials were alternately presented to the participants. A stimulus remained on the screen until a response was given. Participants had 3500 ms to respond. The time interval between the response and

the next stimulus was 1000 ms. The main dependent variables were the median response latencies on repetition and alternation trials.

Smiling Faces (adapted from Rogers & Monsell, 1995; see also Span, Ridderinkhof, & Van der Molen, 2004). Stimuli were schematic faces (man or woman, smiling or unsmiling) that appeared in a 2x2 grid. On each trial a stimulus was presented in the center of one of the four (5x5 cm) quadrants. Participants were required to respond to gender if the face was presented in either one of the top two quadrants, or to the expression of the face if presented in one of the two lower quadrants (blocks 1 and 2, in randomized order; 30 practice trials and 50 experimental trials per block). In the third block, the stimulus moved clockwise through the grid and participants had to respond to gender, if the face occurred in one of the top two quadrants, or to the expression of the face, if it occurred in one of the two lower quadrants (block 3; 90 practice trials, 150 experimental trials). The stimulus remained on the screen until a response was given. Participants had 3500 ms to respond. The time interval between the response and the stimulus of the subsequent trial varied randomly between 900 and 1100 ms (drawn from a uniform distribution). The main dependent variables were the median response latencies on repetition and alternation trials.

6.2.2.3   Inhibition

Stop-signal. In the present version of the Stop-signal task (adapted from Van Boxtel, Van der Molen, Jennings, & Brunia, 2001), participants had to respond as fast as possible to a left or right pointing arrow by a left or right button press. On 25% of the trials, the color of the arrow changed (unpredictably) from green to red, indicating that the response to the arrow stimulus should be inhibited. The time interval between arrow onset and arrow color varied depending on the participant's performance. A dynamic tracking algorithm was used to ensure that stopping approximated 50% correct inhibited responses. The stimulus remained on the screen until a response was given. Participants had 1250 ms to respond. The time interval between the response and the next arrow onset on the subsequent trial varied randomly between 1650 and 2150 ms (drawn from a uniform distribution). There were 50 practice trials, and two blocks of 100 experimental trials. The main dependent variable was the mean stop stimulus reaction time (SSRT), reflecting the latency of the internal response to the stop signal (see Logan, 1994).

Eriksen Flankers (adapted from Ridderinkhof & Van der Molen, 1995). The participant's task was to respond to a left versus right pointing arrow presented at the center of the screen by pressing a left or right response button. The arrow was flanked by four arrows pointing in the same direction (i.e., → → → → → or ← ← ← ← ← ; congruent condition) or by four arrows pointing in the opposite direction (i.e., → → ← → → or ← ← → ← ← ; incongruent condition). The arrow array was presented in a

rectangle, which served as the warning stimulus, as its onset was followed by the arrow array. The time interval between the onset of the warning stimulus and the presentation of the arrow array was 500 ms. The arrow array remained on the screen until a response was given. Participants had 2500 ms to respond. The time interval between the response and the presentation of the warning stimulus of the subsequent trial was fixed at 1000 ms. There were 50 practice trials and 50 experimental trials (i.e., 50 congruent trials and 50 incongruent trials, varied pseudo-randomly). The main dependent variables were the median response latencies on congruent and incongruent trials.

Stroop (adapted from Stroop, 1935). Stimuli were pictures of "smileys" either with a blue or red contour (color task) and either right-side-up or upside-down oriented (orientation task). On the color task, participants had to respond with their dominant hand to the blue or the red contour of the stimulus (pure block 1; 30 practice trials, 85 experimental trials). On the orientation task, participants had to respond to the straight or the upside-down orientation of the stimulus (pure block 2; 30 practice trials, 85 experimental trials). Blocks 1 and 2 were administered in random order and were presented to establish a strong mapping between stimulus (color or orientation) and responses. On a third block (30 practice trials, 100 experimental trials), participants responded to interference trials using the same hand. They responded to one stimulus color (e.g., red), but only when the orientation of stimulus was right-side-up and to the other stimulus color (e.g., blue) but only when the orientation of the stimulus was upside-down. Stimulus color and orientation corresponded on half of the trials. The stimulus remained on the screen until a response was given. Participants had 900 ms to respond. The time interval between the response and the stimulus of the subsequent trial varied randomly between 600 and 800 ms (drawn from a uniform distribution). The main dependent variable was the median response latency on interference trials.

6.2.2.4   Complex EF tasks

Wisconsin Card Sorting Task (WCST). We used a computerized version of the WCST (Somsen, Van der Molen, Jennings, & Van Beek, 2000). Against a light-gray background, four key cards were presented at the top of the screen and were numbered from 1 to 4. The response cards were taken from the original version of the WCST (Grant & Berg, 1948) and were presented one at a time at the bottom of the screen. The task required participants to match the series of response cards with any of four key cards by pressing the number corresponding to that key card. The display remained until a choice was given. Feedback followed the response, and consisted of a "+" sign if the response was correct, or a "-"sign if the response was incorrect.

Response cards could be matched on color (red, green, blue, yellow), shape (triangle, star, cross, circle), or number (1, 2, 3, 4). Once the participant made 10

consecutive correct sorts, the sorting principle changed. The task was terminated either after the participant completed 6 categories (e.g., shape, color, form, color, form, shape), or after the maximum of 128 trials was reached. The order of the sorting principles was randomized, with the constraint that the same sorting principles did not occur consecutively. The test was administered according to the procedure outlined in the Heaton manual (Heaton et al., 1993). The variables of interest were the number of categories achieved, the proportion of perseverative errors (i.e., the total number of errors that occur when a participant is required to switch to another sorting principle, and then persists in responding to the previously correct sorting principle, relative to the number of trials administered, and multiplied by 100), and the proportion of conceptual level responses (i.e., the total number of consecutive correct responses that occur in runs of three or more relative to the number of trials administered, and multiplied by 100).

Tower of London (ToL). The ToL (Shallice, 1982) requires the moving of differently colored balls across three differently sized pegs in order to duplicate a pre-specified target configuration. The smallest peg can hold one ball, the medium-sized two, and the largest peg can hold three balls. Three constraints apply: 1) do not place more than the permitted number of balls on one peg; 2) do not place the balls anywhere other than on a peg; 3) only move one ball at a time. We used a computerized version of the ToL (see for a similar ToL task format: Wiers, Gunning, & Sergeant, 1998). Against a light-gray background, a pegboard with three balls (red, green, blue) positioned on the pegs was presented at the center of the computer screen and a task-specification box (depicting the goal state) in the upper right-hand corner of the computer screen. The balls were repositioned by dragging (with a computer mouse) a ball directly to the position indicated by the task-specification box. The ToL items were taken from Schnirman, Welsh, & Retzlaff (1998). Three difficulty levels were presented: one block of five 4-move trials, one block of five 5-move trials, and one block of five 6-move trials. Performance was scored in terms of the proportion of perfect solutions (i.e., the number of items solved in the minimum number of moves relative to the number of items administered, and multiplied by 100), the mean number of additional moves (i.e., the mean number of moves exceeding the minimum number of moves across each difficulty level), and planning time (i.e., the interval between the occurrence of an item and the first mouse click on one of the balls to be moved).

6.2.3   Procedure

Testing took place in two sessions. In each session, six tasks were administered. The 7-year-olds were tested simultaneously in groups of two. The 11-year-olds, 15-year-olds, and 21-year-olds were tested simultaneously in groups of four. Each test session lasted approximately 1.5 hours. The order of tasks was counterbalanced across participants; the

WCST and ToL were always administered last, considering the inter-individual variation in total time needed to complete the task. There were 3-minute-breaks between tasks, and a 10-minute break following 3 tasks. At the end of the test-session, the 15-year-olds and 21-year-olds completed a paper-and-pencil version of the Raven SPM (Raven et al., 1985); the children completed this task individually in the classroom (with all participants present).

6.2.4   Outlier analysis

If performance was less than 55% correct on one of the conditions in the Shifting and Inhibition tasks (except the Stop-signal task), the results from this particular task were coded missing. The same was done for the Stop-signal task in case the proportion of correct inhibits was lower than 20% or higher than 80%. For all nine EF tasks where RT served as a dependent measure, we performed a two-stage trimming procedure. 1) All incorrect responses, as well as responses that were preceded by an incorrect response, or responses with RTs shorter than 120 ms, or with a latency exceeding the mean by more than 2.5 standard deviations (for each participant and task, separately) were excluded from the RT analyses. This amounted to less than 1.5 % of all trials. 2) Extreme outliers at group-level per condition were identified by SPSS box-plot procedure (SPSS, Inc., 2003). Extreme data (i.e., more than three times the inter-quartile range) were scored as missing. For the Working Memory tasks, we performed the abovementioned step 2 on the accuracy data, because differences between conditions were more prominent for accuracy rather than RT. Accuracy scores were transformed by square root. Failure to complete the first category of the WCST was scored as missing. Only correctly solved items of the ToL were analyzed. Finally, failure to attend the second test session and equipment malfunction resulted in the loss of 4 11-year-olds and 5 15-year-olds. Missing values amounted to 11% of all observations for the 7-year-olds, 6% for the 11-year-olds, 7% for 15-year-olds, and 5 % for the 21-yr olds.

6.3   Results

We performed three sets of analyses. The first set included analyses of variance to assess developmental trajectories for each task. The second set included multi-group confirmatory factor analysis, to assess when the latent components Working Memory, Shifting, and Inhibition reached adult levels. The third set included regression analyses to assess the contribution of the latent factors to the performance on the WSCT and ToL.

6.3.1   Analysis of variance

Proportions correct and median RTs of the EF tasks in the domains of Working Memory, Shifting, and Inhibition are presented in Tables 1 and 2 for each age group separately. Below, we report the accuracy results for the Working Memory tasks and the RT results for the Shifting and Inhibition tasks, because these were the dependent measures for which Condition and Age effects were most prominent. For the RT effects in the Working Memory tasks and the accuracy effects in the Shifting and Inhibition tasks the Condition x Age interactions were either non-significant or in the same direction. Previous research suggested that developmental differences (i.e., comparing children to adults) in task performance can be accounted for by general speeding (e.g., Cerella & Hale, 1994; Kail, 1991; Span et al., 2004). Therefore, in the present study we corrected the RT results for within-group individual differences in basic processing speed. The median RTs obtained from the pure blocks of the Stroop task were taken to provide an estimate of the basic processing speed.

Working Memory. The accuracy scores on the Working Memory tasks were submitted to an ANOVA with Condition (two levels) as within-subjects variable and Age Group (four levels) as between-subject variable. The ANOVA of the Tic Tac Toe task scores resulted in a main effect of Condition, $F(1, 351) = 300.21$, $p < .001$, $MSE = .46$, $\eta_p^2 = .46$; Age Group, $F(3, 351) = 116.86$, $p < .001$, $MSE = .56$, $\eta_p^2 = .50$; and a Condition x Age Group interaction, $F(3, 351) = 44.26$, $p < .001$, $MSE = .46$, $\eta_p^2 = .27$. Post-hoc Bonferroni comparisons for the difference score of the Condition effect showed that the difference between the high and low memory load conditions was larger in 7-year-olds than in 11-year-olds, and larger in 11-year-olds than in 15-year-olds. Fifteen-year-olds did not differ from 21-year-olds.

A similar ANOVA for the Mental Counters resulted in a main effect of Condition, $F(1, 283) = 67.84$, $p < .001$, $MSE = .34$, $\eta_p^2 = .19$; Age Group, $F(3, 283) = 21.31$, $p < .001$, $MSE = .72$, $\eta_p^2 = .18$; and a Condition x Age Group interaction, $F(3, 283) = 6.02$, $p = .001$, $MSE = .34$, $\eta_p^2 = .06$. Again, post-hoc Bonferroni comparisons for the difference score of the Condition effect showed that across blocks the difference between series consisting of 5 stimuli and series consisting of 7 stimuli was larger in 7-year-olds than in 11-year-olds, and larger in 11-year-olds than in 15-year-olds. Fifteen-year-olds did not differ from 21-year-olds.

Finally, the ANOVA for the Running Memory task only yielded a main effect of Age Group, $F(3, 354) = 16.90$, $p < .001$, $MSE = 1.19$, $\eta_p^2 = .13$, showing that performance increased with age. The main effect of Load was not significant, and did not interact with Age Group (F's < 1). Post-hoc Bonferroni comparisons of the accuracy results indicated that accuracy in 7-year-olds was lower than in 11-year-olds, and 11-year-olds did not differ from 15-year-olds (who in turn did not differ from 21-year-olds).

To summarize, on two of the three Working Memory tasks (the Tic Tac Toe task and the Mental Counters task), adult level of performance was not reached until the age

of 15. The Running Memory task showed that accuracy increased with advancing age but in the absence of an interaction with task memory this developmental trend is difficult to interpret.

Shifting. The median RTs on the Shifting tasks were submitted to a MANCOVA with the basic processing speed entered as covariate, Condition (two levels) as within-subjects variable, and Age Group (four levels) as between-subjects variable. The ANCOVA for the Local-Global task resulted in a main effect of Age Group, F (3, 314) = 30.92, p < .001, MSE = 17106.58, η_p^2 = .23; and a Condition x Age Group interaction, F (3, 314) = 7.24, p < .001, MSE = 3330.02, η_p^2 = .07.

The ANCOVA for the Dots-Triangles task resulted in a main effect of Condition, F (1, 315) = 4.37, p = .037, MSE = 20634.45, η_p^2 = .02; Age Group, F (3, 315) = 17.21, p < .001, MSE = 106150.38, η_p^2 = .14; and a Condition x Age Group interaction, F (3, 315) = 10.53, p < .001, MSE = 20634.45, η_p^2 = .09.

The ANCOVA of the scores on the Smiling Faces task also resulted in a main effect of Age Group, F (3, 306) = 11.00, p < .001, MSE = 74719.36, η_p^2 = .10; and a Condition x Age Group interaction, F (3, 306) = 4.95, p = .002, MSE = 21986.26, η_p^2 = .05. Post-hoc Bonferroni comparisons for the shift costs (i.e., the difference in RT between alternation trials and repetition trials in the mixed block) showed that for the Local-Global, the Dots-Triangles, and the Smiling Faces tasks the shift cost was larger in 7-year-olds than in 11-year-olds, and larger in 11-year-olds than in 15-year-olds. However 15-year-olds did not differ from 21-year-olds.

To summarize, for all three Shifting tasks adult level of performance was not reached until the age of 15.

Inhibition. The ANOVA of the Stop-signal task data resulted in a main effect of Age Group, F (3, 283) = 3.94, p = .009, MSE = 4043.96, η_p^2 = .040. Post-hoc Bonferroni comparisons showed that the SSRT was larger for 7-year-olds than for 11-year-olds, larger for 11-year-olds than for 15-year-olds. However, 15-year-olds did not differ from 21-year-olds.

The median RTs on the Eriksen Flankers and the Stroop task were submitted to a MANCOVA with the basic processing speed entered as covariate, Condition (two levels) as within-subjects variable, and Age Group (four levels) as between-subjects variable. The ANCOVA of the Eriksen Flankers task data resulted in a main effect of Condition, F (1, 326) = 9.90, p = .002, MSE = 1306.78, η_p^2 = .03; Age Group, F (3, 326) = 50.70, p < .001, MSE = 12400.09, η_p^2 = .32; and a Condition x Age Group interaction, F (3, 326) = 12.73, p < .001, MSE = 1306.78, η_p^2 = .11. Post-hoc Bonferroni comparisons of the difference score of the Condition effect showed that the difference between congruent and incongruent trials was larger for 7-year-olds than for 11-year-olds, larger for 11-year-olds than for 15-year-olds. Again, 15-year-olds did not differ from 21-year-olds.

EF domain	Task	7-yr-olds	11-yr-olds	15-yr-olds	21-yr-olds
Working Memory	Tic Tac Toe				
	low mem. load	83.4 (1.10)	93.1 (.97)	95.1 (.98)	96.0 (.98)
	high mem. load	51.0 (1.57)	67.7 (1.58)	87.0 (1.68)	90.1 (1.90)
	Mental Counters				
	2 ctrs & ser 5	78.5 (2.30)	84.5 (1.65)	85.0 (1.66)	86.9 (1.68)
	3 ctrs & ser 5	76.2 (2.44)	83.7 (1.83)	85.6 (1.67)	90.4 (1.71)
	2 ctrs & ser 7	61.0 (2.66)	73.4 (2.23)	78.1 (1.94)	85.4 (2.22)
	3 ctrs & ser 7	70.9 (2.86)	73.6 (2.06)	82.4 (2.00)	86.3 (2.04)
	Running Memory				
	block 1	76.6 (1.75)	85.6 (1.48)	86.7 (1.49)	88.9 (1.51)
	block 2	76.7 (2.10)	82.6 (1.64)	87.4 (1.68)	88.2 (1.88)
	block 3	75.3 (1.91)	84.5 (1.65)	88.0 (1.50)	90.6 (1.71)
Shifting	Local-Global				
	repetition	90.3 (.76)	92.9 (.58)	93.5 (.58)	95.1 (.59)
	alternation	88.9 (.75)	92.0 (.77)	93.5 (.77)	95.8 (.78)
	Dots-Triangles				
	repetition	78.5 (1.06)	89.1 (.76)	92.7 (.77)	94.3 (.97)
	alternation	71.4 (1.18)	79.7 (1.07)	84.8 (1.11)	87.8 (1.12)
	Smiling Faces				
	repetition	79.2 (1.25)	85.6 (1.11)	87.8 (1.12)	93.1 (1.35)
	alternation	73.8 (1.20)	79.9 (1.07)	82.8 (1.09)	88.9 (1.13)
Inhibition	Stop-signal				
	% corr. inhibits	46.0 (.95)	48.2 (.69)	50.0 (.85)	47.9 (.83)
	Eriksen Flankers				
	congruent	96.2 (.39)	97.8 (.79)	98.6 (.40)	99.2 (.40)
	incongruent	90.8 (.76)	93.3 (.77)	93.9 (.78)	94.1 (.78)
	Stroop				
	interference	80.8 (.90)	90.8 (.95)	93.1 (.97)	96.6 (.98)

Table 1. Proportions correct (%) for the EF tasks in the domains of Working Memory, Shifting and Inhibition, the control factor Basic Speed. Standard errors are presented between parentheses. NB. ctrs = counters; ser = series; corr. = correct

EF domain	Task	7-yr-olds	11-yr-olds	15-yr-olds	21-yr-olds
Complex	ToL				
	perfect sol. (%)	6.11 (.10)	8.11 (.08)	12.56 (.08)	14.91 (.09)
	add. moves (#)	14.08 (.87)	17.27 (.69)	12.44 (.68)	10.80 (.73)
	plan time (s)	5.78 (.49)	4.20 (.39)	6.22 (.38)	7.55 (.41)
Basic Speed	Stroop				
	pure blocks	97.8 (.20)	99.6 (.20)	99.8 (.20)	99.8 (.20)

Table 1 continued. Proportions correct (%) for the performance measures of the WCST and ToL. Standard errors are presented between parentheses. NB. sol = solutions; add. = additional

Finally, the ANCOVA of the interference RTs on the Stroop task resulted in a main effect of Age Group, F (3, 333) = 6.97, $p < .001$, MSE = 2953.43, η_p^2 = .06. Post-hoc Bonferroni comparisons showed that the interference RT was larger in 7-year-olds than in 11-year-olds, larger in 11-year-olds than in 15-year-olds, and larger in 15-year-olds than in 21-year-olds.

To summarize, on two of the Inhibition tasks (the Stop-signal task and the Eriksen Flankers task), adult level of performance was not reached until the age of 15, whereas for the third task (the Stroop task) performance continued to improve into young-adulthood.

Wisconsin Card Sorting Task. The proportion of conceptual level responses, the number of categories completed, and the proportion of perseverative errors on the WCST were submitted to ANOVAs with Age Group as between-subjects variable. All main effects were significant, F (3, 346) = 13.96, $p < .001$, MSE = 360.48, η_p^2 = .11; F (3, 346) = 23.82, $p < .001$, MSE = .2.90, η_p^2 = .17; and F (3, 346) = 18.46, $p < .001$, MSE = 185.25, η_p^2 = .14, respectively. Post-hoc Bonferroni comparisons showed that the proportion of conceptual level responses was smaller in 7-year-olds than in 11-year-olds, and smaller in 11-year-olds than in 15-year-olds. Fifteen-year-olds, however, did not differ from 21-year-olds. The number of categories was smaller in 7-year-olds than in 11-year-olds, smaller in 11-year-olds than in 15-year-olds, and smaller in 15-year-olds than in 21-year-olds. The proportion of perseverative errors were larger in 7-year-olds than in 11-year-olds, larger in 11-year-olds than in 15-year-olds, and larger in 15-year-olds than in 21-year-olds.

To summarize, for one of the three measures of the WCST (the proportion of conceptual level responses), adult level of performance was not reached until the age of 15, whereas for the other two measures (i.e., the number of categories completed and the

EF domain	Task	7-yr-olds	11-yr-olds	15-yr-olds	21-yr-olds
Shifting	Local-Global				
	repetition	874 (12)	535 (10)	422 (11)	393 (11)
	alternation	1088 (18)	625 (15)	477 (15)	448 (16)
	Dots-Triangles				
	repetition	1164 (23)	810 (18)	623 (18)	588 (20)
	alternation	1677 (40)	1200 (31)	819 (30)	817 (33)
	Smiling Faces				
	repetition	1071 (23)	726 (19)	591 (19)	518 (21)
	alternation	1567 (38)	1088 (31)	818 (31)	703 (33)
Inhibition	Stop-signal				
	Stop-signal RT	289 (8)	238 (7)	217 (7)	201 (8)
	Eriksen Flankers				
	congruent	815 (10)	523 (9)	415 (9)	394 (9)
	incongruent	950 (12)	587 (10)	466 (11)	444 (11)
	Stroop				
	interference	776 (8)	612 (8)	548 (8)	514 (8)
Basic Speed	Stroop				
	pure blocks	579 (6)	424 (6)	356 (6)	341 (6)

Table 2. Median response latencies (ms) for the EF tasks in the domains of Shifting and Inhibition, and for the control factor Basic Speed. Standard errors are presented within brackets

proportion of perseverative errors) performance continued to develop into young-adulthood.

Tower of London. The number of perfect solutions on the ToL, the number of additional moves, and planning time were submitted to ANOVAs with Age Group as between-subjects variable. All main effects were significant, F (3, 348) = 41.02, $p < .001$, MSE = 31.55, η_p^2 = .26; F (3, 348) = 18.21, $p < .001$, MSE = 47.38, η_p^2 = .14; F (3, 348) = 12.17, $p < .001$, MSE = 14.83, η_p^2 = .10. Post-hoc Bonferroni comparisons showed that the number of perfect solutions was smaller in 7-year-olds than in 11-year-olds, smaller in 11-year-olds than in 15-year-olds, and smaller in 15-year-olds than for 21-year-olds. Additionally, the number of additional moves was larger in 7-year-olds than in 11-year-olds, larger in 11-year-olds than in 15-year-olds, while 15-year-olds did not differ from 21-year-olds. Finally, planning time was longer in 7-year-olds than in 11-year-olds,

longer in 11-year-olds than in 15-year-olds. Fifteen-year-olds however did not differ from 21-year-olds.

To summarize, for two of the three measures of the ToL (i.e., the number of additional and planning time), adult level of performance was not reached until the age of 15, whereas for the third measure (i.e., the number of perfect solutions) performance continued to develop into young-adulthood.

6.3.3 Structural equation modeling
6.3.3.1 Model specification

Structural equation modeling (SEM) was used to formulate and test models of EF explicitly in multi-group analyses. The details of the fitting procedure are described in Appendix A. Working Memory was indexed by the mean accuracy on the Tic Tac Toe and of the Mental Counters tasks. The results of the Running Memory task were not included because of the lack of a task load effect. Shifting was indexed by the median RT on shift trials on the Local-Global task, the Dots-Triangles task, and the Smiling Faces. The three manifest Inhibition variables were the median RT on incongruent trials at the Eriksen Flankers task, the mean latency of the SSRT at the Stop-signal task, and the median RT on interference trials at the Stroop task. Basic processing speed, as indexed by median RT of the pure blocks of the Stroop task, was used as a control factor. As the three Shifting tasks and the Eriksen-Flankers and Stroop tasks are RT tasks, these were specified to load on the basic processing speed factor to correct for within-group individual differences in processing speed. Note that the Stop-signal RT task was not specified to load on the basic processing speed factor, because the horse-race model used to estimate SSRT assumes that SSRT and Go RT are independent processes[2] (Logan, 1994).

We first fitted a model to compare EF task performance across groups. We refer to this model as the "partial measurement model"[3], as it involved relating the indicators of Working Memory and Shifting to the common Working Memory and Shifting factors. The indicators of Inhibition were not related to a common Inhibition factor, as the correlations (ranging from -.12 to .35) between the Inhibition tasks were such that a

---

[2] The data obtained in this study are in accordance with the horse-race model that argues that performance on the Stop-signal task reflects a race between activation and Inhibition processes (see Logan, 1994). An ANCOVA on Stop-signal RT, entering Go RT as a covariate yielded a significant main effect of Age Group $F(3,321) = 30.58$, $p < .001$, with Age Group explaining 23% of the variance. As predicted by the horse-race model, it was observed that responses on stop trials that escape Inhibition were executed faster than average go-signal responses in all age groups, $F(1,325) = 825.21$, $p < .001$; mean RT of failed inhibits was 488 ms ($SD = 121$), and mean go-RTs was 570 ms ($SD = 144$).

[3] Task reliabilities were satisfactory. The reliabilities of the tasks (as indexed by indicators of the partial measurement model) were calculated with Cronbach's α (odd-even). Across age groups, the reliabilities ranged from .63 to .91.

common Inhibition factor could not be extracted. Rather than choosing to include one of the Inhibition tasks in the partial measurement model, we decided to include all three tasks as measures of Stop-signal, Eriksen Flankers, and Stroop inhibition, respectively. Thus, the partial measurement model incorporated two latent variables (Working Memory and Shifting), three manifest Inhibition variables, and one control factor.

The model was fitted to the data of all age groups, simultaneously in a multi-group model, to assess whether the indicators of Working Memory and Shifting actually measure the same components across age, and whether observed mean differences are attributable to the common factor mean differences, i.e., we investigated factorial invariance (Meredith, 1993). In fitting the partial measurement model, we introduced various substantive and identifying constraints. We compared the groups by fitting a number of models. First, in order to establish that we measured the same components in all age groups, we defined the factor model M1 within each age group (i.e., we established configural invariance; a path diagram of this model is shown in Figure 1). In this model, the configuration of factor loadings was identical for all age groups, but parameters were free to vary across groups. This model provided a baseline, by which we evaluated more constrained models. Secondly, we established measurement invariance, that is, we constrained the regression coefficients (factor loadings) of the observed indicators of the common factors to be identical in all age groups. (Note that this does not imply that the group means are identical.) This constrained model was coined M2. Finally, we proceeded to assess the developmental trajectories of EF. Thus, we fitted a model, coined M3, in which the 21-year-olds served as a baseline, and we estimated the mean differences of the 7-, 11-, and 15-year-olds relative to the 21-year-olds. With respect to the Working Memory and Shifting variables, model M3 specified strong factorial invariance (Meredith, 1993). This implies that the observed mean differences between the groups with respect to Working Memory and Shifting are a function of the mean differences of the common factors Working Memory and Shifting. From the perspective of linear regression, model M3 specified that across the groups, the regression coefficients and intercepts in the regression of the observed variables on the factors were equal.

To avoid computational problems that may arise when the variances of the variables differ greatly (as is the case here), we rescaled the measurement variables so that the variances of the young-adult group all equaled about 20 (the choice of the value is obviously arbitrary), and we applied the same linear transformation in the other groups. Rescaling the variables does not alter the group differences in any meaningful way, i.e., the group differences in means and covariance structure are retained, albeit on a linearly transformed scale. Subsequently the PRELIS program (Jöreskog & Sörbom, 1999) was
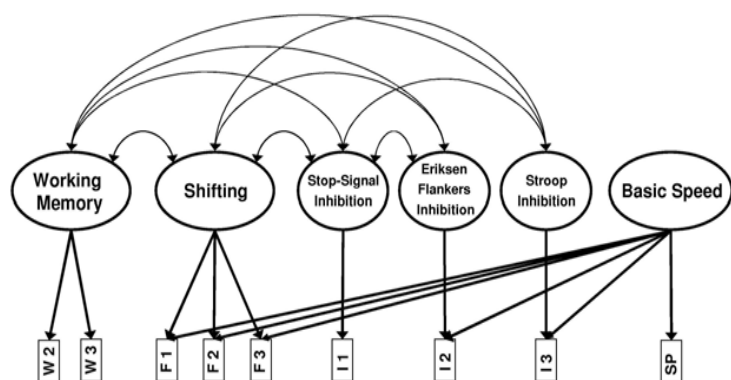
Figure 1. The partial measurement model used for the confirmatory factor analysis. The ellipses represent the latent factors Working Memory and Shifting; the manifest Inhibition variables and the control factor (Basic Speed). The rectangles represent the individual tasks that were chosen to tap the specific EF components. The curved double-headed arrows represent correlations among the latent variables, the Inhibition variables (Stop-signal Inhibition; Eriksen Flanker Inhibition; Stroop Inhibition), and the control factor. Note: W2 = Tic Tac Toe task; W3 = Mental Counters task; S1 = Local – Global task; S2 = Dots – Triangles task; S3 = Smiling Faces task; I1 = Stop-signal task; I2 = Eriksen Flankers task; I3 = Stroop task; SP = Basic Speed

used to transform the variables by calculating normal scores within in each age group. This transformation ensured that skewness and kurtosis of the variables were as close as possible to the values expected under normality (i.e., 0 and 3 respectively). The transformation does not alter the means or the variances of the variables. Because of missingness, we employed raw data likelihood estimation (Schafer & Graham, 2002; Wothke, 2000). That is, rather than fitting the model to the sample summary statistics of the four groups, we fitted the models directly to the data. In doing so, we assumed that the data are missing completely at random (MCAR; see Schafer & Graham, 2002). Judging by the formal test (SPSS Inc., 2003), MCAR appears to hold in the present groups, with the possible exception of the 21-year-olds. Here, however, the value of the χ^2 test statistic was hardly cause for alarm (7-year olds: χ^2 (414) = 406.31, p = .59; 11-year olds: χ^2 (498) = 535.97, p = .11; 15-year olds: χ^2 (473) = 496.63, p = .22; 21-year olds: χ^2 (362) = 419.90, p = .02).

6.3.3.2 Model fitting

Model fitting was done using LISREL (Jöreskog & Sörbom, 1999) and Mx (Neale, Boker, Xie, & Maes, 2003). As we applied raw data likelihood estimation, the usual array of goodness of fit measures is not available. We therefore focused mainly on the χ^2 difference statistics (i.e., loglikelihood ratio tests) to investigate the measurement model and measurement invariance.

First, to establish configural invariance, we fitted the factor model M1. This model served as the baseline model. We found the fit of M1 to be acceptable. The formal goodness of fit, as indexed by the χ^2, equaled 139.34 (df = 67). Next, we introduced (and established) measurement invariance (M2). Following the imposition of the M2 equality constraints, and compared to M1, we observed a χ^2 difference of 17.85 (df = 21, n.s.). The percentages of variances of the data of the five tasks of Working Memory and Shifting explained by the common factors of Working Memory and Shifting are presented in Table 3. To determine the adequacy of the models M1 and M2, we inspected the range of the standardized residuals. (These are a function of the difference between the unconstrained covariance matrix and the expected covariance matrix under the specified model (both matrices are estimated using raw data likelihood estimation). In a well fitting model, they are approximately standard normally distributed, and thus should vary between about –3 and 3. In the three oldest age groups, the standardized residuals were acceptable (between –3.6 and 2.8); for the youngest age group only four out of the 45 to be estimated standardized residuals were greater than –3 or greater than 3. These standardized residuals were all related to Mental Counters task performance.

Factor	Task	7-yr-olds	11-yr-olds	15-yr-olds	21-yr-olds
Working	Tic Tac Toe	46	7	20	47
Memory	Mental Counters	75	12	25	51
Shifting	Local-Global	28	50	64	67
	Dots-Triangles	17	41	39	34
	Smiling Faces	29	45	41	60
Basic Speed	Stroop pure blocks	83	77	86	76

Table 3. Percentage of variance explained by the latent factors Working Memory and Shifting, and by the control factor Basic Speed

Finally, we performed a 4-group confirmatory factor analysis. We therefore introduced the mean structure into the model (M3), and compared to M2, we observed a χ^2 difference of 51.45 (df = 9; p < .001). Inspection of the standardized residuals of the

covariance structure (between –3.6 and 3.1 in the oldest three age groups), and the standardized residuals for the means (between –1.7 and 2.8 in all age groups) suggest that M3 is acceptable. The correlations between the common factors of Working Memory and Shifting, and the three Inhibition tasks are presented in Table 4. Note that the hypothesis of a single common factor in the present normal populations is unlikely, in the light of the generally low correlations between the common factors.

	7-yr-olds	11-yr-olds	15-yr-olds	21-yr-olds
Working Memory – Shifting	.40	.04	-.08	-.54*
Working Memory – Stop-signal	.09	-.27	-.42*	-.55*
Working Memory – Eriksen Flankers	-.03	-.11	.17	-.21
Working Memory – Stroop	-.14	-.17	.36	-.20
Shifting – Stop-signal	-.85*	.16	.09	.27
Shifting – Eriksen Flankers	.60	.34*	.41*	.35*
Shifting – Stroop	-.35	.02	-.10	.26*
Stop-signal – Eriksen Flankers	.12	.26*	.09	.26*
Stop-signal – Stroop	-.07	.30*	-.12	.19
Eriksen Flankers – Stroop	-	-	-	-

*Table 4. Correlations of the latent factors Working Memory and Shifting and the manifest Inhibition measures for each of the four age groups. * = significant. Note: -: fixed to zero*

	α^*1	α^*2	α^*3
Working Memory	-4.83 (.72)*	-1.57 (.28)*	-.49 (.19)*
Shifting	1.76 (.71)*	1.17 (.33)*	.09 (.19)
Stop-signal Inhibition	1.54 (.26)*	.67 (.17)*	.29 (.17)
Eriksen Flankers Inhibition	6.81 (.99)*	1.97 (.40)*	.23 (.21)
Stroop Inhibition	-1.09 (.74)	-.15 (.31)	.28 (.17)
Basic Speed	7.92 (.90)*	2.88 (.37)*	.62 (.18)*

*Table 5. Maximum likelihood estimates and standard errors (between parentheses) of differences in latent factor means of Working Memory and Shifting, the three manifest Inhibition variables and the control factor Basic Speed. Note: * significant difference relative to 21-year-olds*
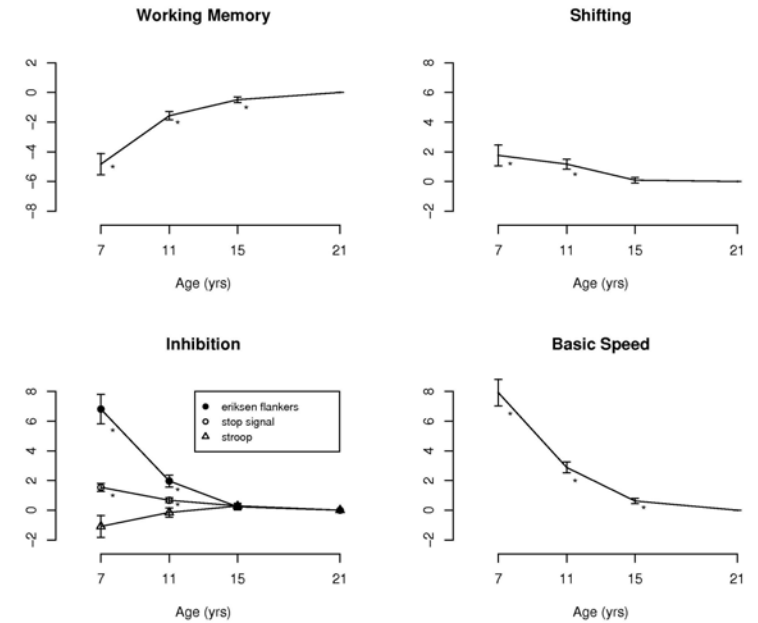
*Figure 2. Maximum likelihood estimates and their standard errors of differences in common factor means of Working Memory and Shifting (i.e., the observed mean differences between the groups as a function of the mean differences of the common factors), the three manifest Inhibition variables and the control factor Basic Speed; *: significant difference relative to 21-year-olds*

The maximum likelihood estimates and standard errors of the common factor mean differences (i.e., in the notation of the appendix, the parameters in the vectors α^*1, α^*2, and α^*3) are presented in Table 5. As explained in Appendix A, these parameters represent differences in common factor means relative to the young-adult group. In other words, the observed mean differences between the groups with respect to Working Memory and Shifting are a function of the mean differences of the common factors Working Memory and Shifting. The standard errors of the estimates indicated that 7-year-olds and 11-year-olds differed from 21-year-olds with respect to Working Memory, Shifting, the Stop-signal task, the Eriksen Flankers task and basic processing speed.

Fifteen-year-olds differed from 21-year-olds with respect to Working Memory and basic processing speed. These developmental patterns are apparent in Figure 2 that provides plots of the decomposition of means of each EF relative to the young-adult group.

6.3.3.3   Predicting performance on the WCST and the ToL

In a number of separate analyses the complex EF tasks (i.e., the ToL and WCST tasks) were regressed on the latent variables in the partial measurement model. Again we performed 4-group analyses, as we wished to assess possible age-related changes in the regression equations. The analyses were done in three steps. The details of the regression analyses are presented in Appendix B. First, we tested the hypothesis that the regression coefficients differed between age groups (Model A). Second, we tested the hypothesis that the regression coefficients were equal across the groups (Model B). Note that across-group equality of the regression coefficients does not necessarily imply that the amount of explained variance is equal over the age groups. Third, we investigated whether the regression coefficients were significantly greater than zero (Model C).

The results of the regression analyses of the WCST measures indicated that for the number of categories achieved and the proportion of conceptual level responses, Model C did not fit significantly worse than Models B and A. This finding suggests that the common factors Working Memory and Shifting and the manifest Inhibition variables do not contribute substantially to the number of categories and in the proportion of conceptual level responses of the WCST. In the analysis of the proportion of perseverative errors, we found that the difference between Models A and B was not significant (χ^2 difference = 17.85; df = 15), but the difference between Models C and B were (χ^2 difference = 15.30; df = 5, p <. 010), indicating that the regression coefficients were equal across groups, but not zero. The upper panel of Table 6 presents the estimates of the regression coefficients with 95% confidence intervals. It can be seen that in all age groups Working Memory related strongly to the proportion of perseverative errors. Thus, perseveration behavior on the WCST appeared to be strongly related to Working Memory capacity.

For the mean number of additional moves and planning time on the ToL, we failed to fit a reasonable model. Regression analysis of the proportion of perfect solutions indicated that the difference between Models A and B was not significant (χ^2 difference = 16.60; df = 15). The difference between Models C and A was significant (χ^2 difference = 21.47; df = 5, p < .001), indicating that the regression coefficients differed across groups and do not equal zero. Calculating the 95% confidence intervals of the regression coefficients in the youngest three age groups proved problematic. In the lower panel of Table 6, it is shown that Stroop performance was a significant predictor of the proportion of perfect solutions. Thus, in 21-year-olds, the ability to suppress a pre-potent response,

as on the Stroop task, seems to play an important role in the efficiency of solving ToL items.

6.4   Discussion

In this study, we examined the developmental trajectories of three frequently postulated EF components, Working Memory, Shifting and Inhibition, in relation to performance on standard neuropsychological EF tasks, the WCST and the ToL. In so doing we adopted both standard analyses and multi-group latent variable modeling. The latter enabled us to model the structure of the underlying latent variables across age groups, thus going beyond the usual analysis of observed developmental trends in the performance on the task battery.

6.4.1   Analysis of manifest task performance

First, we analyzed the data using conventional analyses of variance. The developmental trajectories for Working Memory, Shifting and Inhibition that emerged from the task battery were consistent with several developmental studies indicating that developmental trends of different EF component processes are distinct (e.g., Brocki & Bohlin, 2004; Klenberg et al., 2001; Lehto, 2004; Luciana et al., 2005; Luna et al., 2004; Welsh et al., 1991). The correction for within-group individual differences in basic processing speed did not remove the age related trends in task performance. This finding is important, because it indicates that the developmental trends in EF task performance cannot be explained in terms of a global change in basic processing speed (e.g., Cerella & Hale, 1994; Kail, 1991; Span et al., 2004).

Developmental changes in Working Memory are well-established, both at a behavioral level (e.g., Gathercole et al., 2004; Luciana et al., 2005; Luna et al., 2004), and at the level of brain functioning. Neuroimaging studies have shown that development of Working Memory co-occurs with the functional maturation of lateral PFC (e.g., Klingberg, Forssberg, & Westerberg, 2002; Kwon, Reiss, & Menon, 2002). In the current study, Working Memory was tapped using three different tasks, and, consistent with prior studies, we found that adult level of performance was not reached before the age of 12 (see Case, Kurland, & Goldberg, 1982; Cohen & Heath, 1990; DeLuca et al., 2003; Gathercole, 1999; Luciana et al., 2005; Luna et al., 2004; Siegel, 1994). The performance on one of the Working Memory tasks (i.e., the Running Memory task), however, failed to reveal a significant effect of task load. The accuracy of performance increased with age but this age-related trend did not interact with task load. The lack of a task load effect was unexpected and, at this point, we do not have a ready explanation for its absence. Possibly, the lack of a task load effect is due to an interference confound. That is, task

load increases with longer stimulus pair series but potential interference between stimulus pairs decreases, as the number of possible stimulus pairs increases with series length. Unfortunately, there is not sufficient data to test this interpretation.

Developmental studies on task shifting are still scarce, but studies to date have demonstrated that the cost of shifting between tasks decreases with age (e.g., Cepeda et al., 2001). In the current study, Shifting was assessed using three different tasks that required shifting back and forth between tasks. All three tasks showed that shift costs (i.e., the difference in RT and accuracy on shift trials relative to repeat trials) decreased with age until the age of 15. Thus, the ability to shift task sets does not reach young-adult level of performance until adolescence (Cepeda et al., 2001; Crone, Bunge et al., 2006; Huizinga & Van der Molen, 2005a). The age-related change in the ability to shift task sets has been interpreted in terms of two prevailing hypotheses of shift costs. First, it has been argued that the larger shift costs in children reflect immature levels of executive control (Cepeda et al., 2001; Zelazo, Craik, & Booth, 2004). In contrast, Crone, Bunge et al. (2006) interpreted larger shift costs in children in terms of memory retrieval of S-R links. Systematic manipulation within the context of the task-shift paradigm should reveal the basic mechanism(s) underlying developmental change in Shifting ability (e.g., Huizinga & Van der Molen, 2005a).

Inhibition is the most studied measure of EF in development (e.g., Bjorklund & Harnishfeger, 1990; Dempster, 1992; Nigg, 2000). The current focus was on the inhibition of motor responses, indexed by using three different tasks. These tasks yielded distinct developmental trends. The performance on the Eriksen Flankers task and the Stop-signal task improved rapidly until the age of 11, and at this age performance did not differ from either 15-year-olds or 21-year-olds. This developmental trend is consistent with previous studies using the Eriksen Flankers task (e.g., Ridderinkhof & Van der Molen, 1995), and the Stop-signal task (e.g., Van den Wildenberg & Van der Molen, 2004; Williams et al., 1999). In contrast, the current version of the Stroop task yielded only a weak developmental trend from early childhood into adulthood. Previous studies using a child-friendly version of the Stroop task showed a rapid decline in the suppression of pre-potent responses during early childhood (e.g., Wright, Waterman, Prescott, & Murdoch-Eaton, 2003). The current task revealed an early and strong decline in error rate, but not in basic processing speed. The current developmental trends in performance on the Eriksen Flankers task and the Stop-signal task can be interpreted in the light of developmental theories that emphasize the role of Inhibition in cognitive development (e.g., Bjorklund & Harnishfeger, 1990; Dempster, 1992; for a review see: Van der Molen, 2000). These theories assume that the ability to resist interference and inhibit competing responses defines a major developmental dimension promoting performance on a wide array of cognitive tasks (see also Kipp, 2005).

The two standard neuropsychological tasks, the WCST and the ToL, yielded developmental trends that are consistent with the literature. For the WCST, the proportions of perseverative errors and of conceptual level reached young-adult level in 11-year-olds, whereas the number of categories completed increased into young-adulthood (Chelune & Baer, 1986; Welsh et al., 1991). For the ToL, planning time and the number of additional moves reached young-adult level in 11-year-olds, whereas the proportion of perfect solutions increased into young-adulthood (e.g., Anderson et al., 2005; Baker et al., 2001; Lehto, 2004).

Taken together, the performance analyses revealed that Working Memory, Shifting and Inhibition reached adult levels between 11 and 15 years, whereas some aspects of inhibitory control (Stroop task) did not reach adult levels until after the age of 15. The performance on the WCST and the ToL reached adult levels between 11 and 21 years.

6.4.2 Analysis of latent variables

A major aim of the current study was to assess the developmental change in the latent components Working Memory, Shifting, and Inhibition. To address this question, we adopted the SEM approach advanced by Miyake et al. (2000), who established that these EF components are distinct, but moderately correlated, at the latent level. Following this study, we employed multi-group structural equation modeling to evaluate an a priori specified measurement model of EF that included EF components as latent factors underlying age-related changes in EF task performance.

We had to depart from the Miyake et al. approach in that our measurement model included two latent variables, Working Memory and Shifting, three manifest Inhibition variables, and one control factor (basic processing speed). The reason for separating the Inhibition variable into three manifest Inhibition variables was based on preliminary analyses, which revealed low and, in some cases, negative correlations between Inhibition measures. Thus, these measures were retained as separate variables rather than aggregating them in a single common factor (for similar results see Van der Sluis, De Jong, & Van der Leij, 2005). Although several authors have argued that 'inhibition' comes in many varieties depending on the exact paradigm that is used (e.g., Dempster, 1993; Harnishfeger, 1996; Kipp, 2005; Logan, 1994; Nigg, 2000), the current observation of disparate developmental trends in the ability to inhibit speeded motor responses was unexpected. Three paradigms were selected to provide indicators of the ability to inhibit motor responses. In the Eriksen Flankers task, subjects are required to execute a speeded motor response to the central stimulus of the stimulus array and to inhibit the competing response that is activated by the flanker stimuli on incongruent trials. Similarly, in the Stroop task, subjects are required to respond as quickly and accurately as possible to the

color of the stimulus, but to inhibit the activated response when the color and orientation of the stimulus does not match on interference trials. Finally, in the Stop-signal task, subjects have to execute a speeded response to the go stimulus, but must inhibit this response on stop-signal trials. An important similarity is that all three paradigms require a form of inhibition, which concerns the control of competing responses. In this regard, the current finding of low correlations between inhibition measures is difficult to explain. Possibly, subtle differences in task format may have contributed to the dissimilarities across measures. That is, the Eriksen Flankers task is a two-choice task requiring a speeded response on all trials, whereas the Stroop task and the Stop-signal task require a speeded response on some trials, but inhibition of response on other trials. In addition, the Stroop task differs from the Stop-signal task in that a conjunction of stimulus features (color and orientation) determines whether or not a response has to be executed in the Stroop task, whereas the instruction to execute or inhibit a motor response is provided by different signals in the Stop-signal task.

The SEM model presented in Figure 1 provided an acceptable description of EF task performance. The SEM analyses indicated that the latent factors Working Memory and Shifting, and the manifest Inhibition variables are clearly separable, as their intercorrelations are quite modest. The correlations among the latent and manifest variables constituting the measurement model were found to be rather low. This result provides additional support for the non-unitary nature of EF (Fisk & Sharp, 2004; Miyake et al., 2000; Stuss et al., 1995). Most importantly, however, given the present aims, the current study added a developmental dimension to the EF model fitting. The model presented in Figure 1 provides an adequate description of EF task performance in each age group. The analyses revealed developmental trajectories that, by and large, are in accord with the ANOVA results. Seven-year-olds and 11-year-olds differed from 21-year-olds in Working Memory and Shifting, and on two of the Inhibition tasks (i.e., the Stop-signal and Eriksen Flankers tasks, but not the Stroop task). Finally, the child groups differed from 21-year-olds in basic processing speed. Fifteen-year-olds differed from 21-year-olds with respect to Working Memory and basic processing speed. The overall pattern of results suggests that Shifting and performance on the Stop-signal and Eriksen Flankers tasks have reached mature levels by adolescence, while Working Memory and basic processing speed follow a more protracted course of development into young-adulthood. Thus, although analysis of variance of the separate Working Memory tasks suggests that 15-year-olds do not differ from 21-year-olds, when examining Working Memory at a latent level there is still maturation after the age of 15.

The present modeling approach has two important advantages over the standard analysis of variance approach. First the analyses of trends were conducted at the level of the latent theoretical variables, in the case of Shifting and Working Memory (but sadly not in the case of Inhibition). This means that the problem of task impurity is at least

somewhat reduced. Equally important is the fact that the analyses of trends were conducted subject to strong factorial invariance (see Appendix A). The restrictions associated with this model go a long way to ensuring that we are actually measuring the same hypothetical constructs across different age groups. The standard analysis of variance is based on the tacit assumption that the latent variables are identical across the age groups. However, it remains uncertain whether a given mean age difference is actually attributable to development in the same latent variable. By establishing strong factorial invariance first, we can interpret the observed mean differences in terms of the latent mean differences in the latent variables that the tasks purport to measure.

When regressing the WCST performance (as indexed by perseveration errors) on the latent variables in partial measurement model, we found that in all groups Working Memory related significantly to the proportion of perseverative errors. Moreover, the regression coefficients were equal across groups. These results differ from those of Miyake et al. (2000), who showed that Shifting, rather than Working Memory predicted perseverative errors on the WCST. In earlier studies, it was found that patients who have damage to dorsolateral PFC experience difficulties with both Working Memory (e.g., Bechara, Damasio, Tranel, & Anderson, 1998) and the WCST (e.g., Barceló & Knight, 2002). Thus, it is likely that Working Memory is an important contributor to WCST performance as indexed by perseveration errors. The differences in findings between this study and the study by Miyake et al. (2000) may be associated with differences in task instructions. That is, in the current study we did not give instructions about the switching of sorting rules, as required by the Heaton et al. (1993) manual, whereas Miyake et al. (2000) informed participants that sorting principles would switch from time to time. Recently, Stuss et al. (2000) in a patient study showed that test administration (e.g., in advance verbalizing the sorting rules) played a role in the sensitivity of the WCST.

Finally, when regressing the proportion of perfect solutions on the ToL on the latent variables in the partial measurement model, we found that in 21-year-olds the proportion of perfect solutions on the ToL was significantly predicted by Stroop inhibition. The χ^2 difference indicated that the regression coefficients differed across groups. However, calculating the 95% confidence intervals of the regression coefficients in the youngest three age groups proved computationally problematic, which may be due to missingness (see below). The finding that Stroop performance plays a role in ToL performance is largely consistent with Miyake et al. (2000), who found that Inhibition predicted ToH performance, given "that no specific instructions for strategies were given and many people are likely to use a perceptual strategy to solve the ToH" (cf. Miyake et al., 2000, p.87). The ability to suppress a pre-potent response (as indicated by Stroop performance) has been reported earlier to be a significant predictor of ToL performance in a young-adult sample (e.g., Welsh, Satterlee-Cartmell, & Stine, 1999). The finding that Inhibition (as measured by Stroop performance) did not predict ToL performance in the

three child groups suggests that they may have used a different strategy to solve the task than adults did. This should be a focus of research for future studies.

One limitation of the current study concerns missingness. Missing data were retained as missing, and, in contrast to Miyake et al. (2000), were not imputed. Missingness amounted to a loss of 11% of the 7-year-olds data, and about 5% in the older age groups (Miyake et al., 2000 reported 2.2% of the data was affected by their trimming procedures). The data was found missing completely at random, allowing us to fit the models directly to the data using raw data likelihood (rather than to the summary statistics; see Schafer & Graham, 2002). This method of estimation does not involve imputation or any arbitrary treatment of missingness, such as pair-wise or list-wise deletion. Thus, from the point of view of estimation, the missingness does not pose a great problem. However, the missingness is known to affect the power adversely (Dolan, Van der Sluis, & Grasman, 2005). With regard to the prediction of performance on the complex tasks, we unexpectedly found that the regression coefficients in all age groups proved to equal zero, when regressing the proportion of conceptual level responses and the number of categories achieved on the WCST on the common factors in the partial measurement model. In addition, for the mean number of additional moves and planning time on the ToL we failed to fit a reasonable model. This may at least in part be due to the missingness. Certainly, the loss of statistical power will have contributed to the failure to detect significant regression on most measures of WCST and ToL performance (Dolan et al., 2005).

A second limitation of the current study refers to the failure to incorporate the Running Memory task as indicator of the common factor of Working Memory. Thus, the Working Memory factor was defined by only two indicators. Generally, it is desirable to have at least three indicators to define a common factor (e.g., Kline, 1998). However, the identification of the Working Memory factor is still statistically feasible given the presence of other (correlated) common factors. That is, while a two indicator common factor model is not identified in isolation, it is identified in the presence of other common factors, provided the two indicator common factor is correlated with the other common factors, as is the case here. Substantively, this common factor can thus still be interpreted as representing the common influence underlying the two indicators of Working Memory.

6.4.3   Conclusion

The results of the current study provide support for the non-unitary, multi-faceted nature of EF. Two latent factors "Working Memory" and "Shifting" were identified that correlated only moderately. This finding is consistent with other studies taking a latent variables approach and showing separable EF components (e.g., Fisk & Sharp, 2004;

Lehto et al., 2003; Miyake et al., 2000). Importantly, the current study showed different developmental trends in Working Memory and Shifting. Working Memory continued to develop into young-adulthood whereas Shifting attained mature levels during adolescence. The finding of different developmental trends is in accordance with earlier literature showing that Working Memory and Shifting mature at different rates (Cepeda et al., 2001; Crone, Bunge et al., 2006; Gathercole et al., 2004; Huizinga & Van der Molen, 2005a; Luciana et al., 2005; Luna et al., 2004). The measurement model that was fitted to the data also included three manifest Inhibition variables that could not be incorporated into a single latent factor. In spite of the conceptual similarity (i.e., all three paradigms required the inhibition of a motor response), the measures derived from the Eriksen Flankers, Stop-signal, and Stroop paradigms showed low or negative correlations. The ability to inhibit the response to the Go stimulus on Stop-signal trials of the Stop-signal task, and the ability to control response competition on incongruent trials of the Eriksen Flankers task improved rapidly to the age of 11. These findings are consistent with the results of previous studies using these inhibition paradigms (Bédard et al., 2002; Ridderinkhof & Van der Molen, 1995; Van den Wildenberg & Van der Molen, 2004; Williams et al., 1999). On the Stroop task, the speed of responding on interference trials revealed only little developmental change. However, response accuracy increased rapidly during childhood, and continued to develop into young adulthood.

In sum, the current findings correspond well with previous studies showing developmental improvement in the performance on the WCST and the ToL. Adult levels on the WCST and the ToL were attained between 11 and 15 years, although the proportion of perfect solutions on the ToL increased into young-adulthood (Anderson et al., 2005; Baker et al., 2001; Chelune & Baer, 1986; Lehto, 2004; Welsh et al., 1991). In all groups Working Memory related equally and significantly to the proportion of perseverative errors. In addition, the current findings are in accord with previous findings demonstrating developmental improvements in EF component processes and recent findings emerging from brain imaging studies of cognitive development (e.g., Bunge et al., 2002; Casey et al., 1997; Crone, Donohue, Wendelken, Honomichl, & Bunge, 2006; Durston et al., 2002; Klingberg et al., 2002; Kwon et al., 2002; Luna & Sweeney, 2004). Neurophysiological studies showing that the anatomical development of PFC areas only reaches maturity in young-adulthood support this notion (e.g., Caviness, Kennedy, Richelme, Rademacher, & Filipek, 1996; Chugani, Phelps, & Mazziotta, 1987; Huttenlocher, 1979; Sowell et al., 2004; Yakovlev & Lecours, 1967; for a review see: Casey, Tottenham, Liston, & Durston, 2005). In future research, it would be of interest to evaluate the biological plausibility of EF component processes and their development by examining anatomically and functionally separate systems within PFC in concert with a latent variables approach such as presented in this study.

6.5 Appendix A

The partial measurement model incorporated two latent variables, three manifest Inhibition variables, and one control factor (i.e., basic speed). The two latent variables represented the EF components Working Memory and Shifting. The mean accuracy of the Tic Tac Toe task, and the Mental Counters task (denoted T_1 and T_2) were the indicators of Working Memory; the median RT on shift trials on the Local-Global task, the Dots-Triangles task, and the Smiling Faces task (denoted T_3 through T_5) were the indicators of Shifting. The three observed Inhibition tasks were the Eriksen Flankers task (median RT on incongruent trials), the Stop-signal task (mean latencies of the SSRT), and the Stroop task (median RT on interference trials; denoted T_6 through T_8). The control factor represented basic processing speed, as indexed by median RT of the pure blocks of the Stroop task (denoted T_9). The three Shifting tasks and the Eriksen Flankers and Stroop tasks were specified to load on the basic processing speed factor to correct for within-group individual differences in basic RT.

The partial measurement model was fitted in a number of steps. First, we defined the factor model within each group by specifying the following regressions:

First, we considered the Basic Speed factor:

$$T_{9i} \quad = \quad \tau_{9i} + \lambda_{9,6i} * BS_i + \varepsilon_{4i},$$

Note that the index on the factor loading λ refers to the position of the matrix of factor loadings shown below.

Secondly, we considered the Working-Memory (WM) factor (Tic Tac Toe task, Mental Counters task):

$$T_{1i} \quad = \quad \tau_{1i} + \lambda_{11i} * WM_i + \varepsilon_{1i}$$
$$T_{2i} \quad = \quad \tau_{2i} + \lambda_{12i} * WM_i + \varepsilon_{2i}$$

Thirdly, we considered the Shifting (S) factor (Local-Global task, the Dots-Triangles task, and the Smiling Faces task):

$$T_{3i} \quad = \quad \tau_{3i} + \lambda_{32i} * S_i + \lambda_{36i} * BS_i + \varepsilon_{3i}$$
$$T_{4i} \quad = \quad \tau_{4i} + \lambda_{42i} * S_i + \lambda_{46i} * BS_i + \varepsilon_{4i}$$
$$T_{5i} \quad = \quad \tau_{5i} + \lambda_{52i} * S_i + \lambda_{56i} * BS_i + \varepsilon_{5i},$$

Finally, we considered the three Inhibition variables (Eriksen Flankers (EFL) task, Stop-signal (SS) task, Stroop (STR) task):

$$T_{6i} \quad = \quad \tau_{6i} + \lambda_{63i} * EFL_i$$
$$T_{7i} \quad = \quad \tau_{7i} + \lambda_{74i} * SS_i + \lambda_{76i} * BS_i$$
$$T_{8i} \quad = \quad \tau_{8i} + \lambda_{85i} * STR_i + \lambda_{86i} * BS_i$$

In these regression equations, τ represents the intercept, λ the regression coefficient, and ε is a residual term. The subscript i denotes age group (i=1, 2, 3, 4, where 1 = 7-year-olds, 2 = 11-year-olds, 3 = 15-year-olds, 4 = 21-year-olds). It is convenient to express these equations in a matrix expression:

$$\mathbf{T}_i \quad = \quad \boldsymbol{\tau}_i + \boldsymbol{\Lambda}_i \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i, \tag{1}$$

where $\mathbf{T}_i$ is the random vector of the 9 observed variables, $\boldsymbol{\tau}_i$ is the vector of intercepts, and $\boldsymbol{\varepsilon}_i$, is the random vector of mutually uncorrelated residuals. The random vector $\boldsymbol{\eta}_i$ contains the common factors, WM, S, and BS, and the Inhibition variables EFL, SS, STR. The 9 x 6 matrix $\boldsymbol{\Lambda}_i$ contains the regression coefficients:

$$\boldsymbol{\Lambda}_i \quad = \quad \begin{matrix} \lambda_{11} & 0 & 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{32} & 0 & 0 & 0 & \lambda_{36} \\ 0 & \lambda_{42} & 0 & 0 & 0 & \lambda_{46} \\ 0 & \lambda_{52} & 0 & 0 & 0 & \lambda_{56} \\ 0 & 0 & \lambda_{63} & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{74} & 0 & \lambda_{76} \\ 0 & 0 & 0 & 0 & \lambda_{85} & \lambda_{86} \\ 0 & 0 & 0 & 0 & 0 & \lambda_{96} \end{matrix}$$

The residuals were constrained to be mutually uncorrelated, so the covariance matrix of the residuals was diagonal. We denote this covariance matrix $\boldsymbol{\Theta}_i$:

$$\text{diag}(\boldsymbol{\Theta}_i) = [\sigma^2_{\varepsilon 1} \ \sigma^2_{\varepsilon 2} \ \sigma^2_{\varepsilon 3} \ \sigma^2_{\varepsilon 4} \ \sigma^2_{\varepsilon 5} \ 0 \ 0 \ 0 \ \sigma^2_{\varepsilon 9}],$$

where σ^2_{ε} denotes the variance of the residuals. Note that we did not introduce residual variances for the Inhibition variables EFL, SS, STR, as these were not treated as

indicators of Inhibition, but rather simply as indicators of themselves. Finally, we considered the covariance matrix of $\boldsymbol{\eta}_i$, denoted $\boldsymbol{\Psi}_i$.

$\boldsymbol{\Psi}_i=$

σ^2_{WM}					
$\sigma_{S,WM}$	σ^2_{Sh}				
$\sigma_{EFL,WM}$	$\sigma_{EFL,S}$	σ^2_{EFL}			
$\sigma_{SS,WM}$	$\sigma_{SS,S}$	$\sigma^2_{SS,EFL}$	σ^2_{SS}		
$\sigma_{STR,WM}$	$\sigma_{STR,S}$	$\sigma^2_{STR,EFL}$	$\sigma^2_{STR,SS}$	σ^2_{STR}	
0	0	0	0	0	σ^2_{BS}

Note that $\sigma^2_{STR,SS}$ was fixed to zero, because it proved impossible to fit the model in all age groups when $\sigma^2_{STR,SS}$ was estimated freely. Assuming $\boldsymbol{\eta}_i$ and $\boldsymbol{\epsilon}_i$ are uncorrelated, we expressed the expected covariance matrix in group i, $\boldsymbol{\Sigma}_i$, as follows:

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}_i \boldsymbol{\Psi}_i \boldsymbol{\Lambda}_i^t + \boldsymbol{\Theta}_i. \tag{2}$$

An important aim of the present SEM was to incorporate the means, as this allowed us to study group differences with respect to the common factors, rather than with respect to the observed variables, as in the MANCOVA. To this end, we specified the following expressions for the means, denoted $\boldsymbol{\mu}_i$, in matrix notation:

$$\boldsymbol{\mu}_i = \boldsymbol{\tau}_i + \boldsymbol{\Lambda}_i \boldsymbol{\alpha}_i \tag{3}$$

where the 6 dimensional vector $\boldsymbol{\alpha}_i$ equaled $E[\boldsymbol{\eta}_i]$, i.e., the vector of means of the latent variables $\boldsymbol{\eta}_i$, the 9 dimension vector $\boldsymbol{\tau}_i$ contains the intercepts.

In fitting this model, we introduced various substantive and identifying constraints. We compared the groups by fitting a number of models. In the first model (M1), equality constraints were not imposed. In this model, the configuration of factor loadings was identical over the groups, but no parameters were constrained to be equal over the groups. This model provided a baseline, by which we may evaluate other more constrained models:

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}_i \boldsymbol{\Psi}_i \boldsymbol{\Lambda}_i^t + \boldsymbol{\Theta}_i \quad (i=1\dots4)$$
$$\boldsymbol{\mu}_i = \boldsymbol{\tau}_i \quad (i=1\dots4)$$

In fitting this model, we fixed the variance of $\boldsymbol{\eta}_i$ to equal 1 in each group (i.e., scaling constraints) and we estimated the regression coefficients in $\boldsymbol{\Lambda}_i$. In addition, we fixed the means $\boldsymbol{\alpha}_i$ to equal zero, so that there was a one-to-one relationship between the observed means and the intercepts $\boldsymbol{\tau}_i$. Subsequently, we constrained the regression coefficients to be equal in model 2 (M2):

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}\boldsymbol{\Psi}_i\boldsymbol{\Lambda}^t + \boldsymbol{\Theta}_i \quad (i=1\dots4)$$
$$\boldsymbol{\mu}_i = \boldsymbol{\tau}_i \quad (i=1\dots4)$$

Given this constraint, we retained the standardization of $\boldsymbol{\eta}$ in group 4, but estimated the variances of $\boldsymbol{\eta}$ freely in the other 3 groups, i.e., $\boldsymbol{\Psi}_4$ is a correlation matrix, while $\boldsymbol{\Psi}_1$, $\boldsymbol{\Psi}_2$, and $\boldsymbol{\Psi}_3$ are covariance matrices. This expressed the fact that, across age, the variance of the variables of interest may vary. In model M3, we introduced the means. To this end we fitted the following model:

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}\boldsymbol{\Psi}_i\boldsymbol{\Lambda}^t + \boldsymbol{\Theta}_i.$$
$$\boldsymbol{\mu}_i = \boldsymbol{\tau} + \boldsymbol{\Lambda}_i \boldsymbol{\alpha}_i \quad (i=1\dots3)$$
$$\boldsymbol{\mu}_4 = \boldsymbol{\tau}$$

Note that we equated the means and the intercepts in group 4 ($\boldsymbol{\mu}_4 = \boldsymbol{\tau}$), and estimated the parameter vector $\boldsymbol{\alpha}$ in groups 1 to 3. In this parameterization, group 4 served as a reference group and the parameter vectors $\boldsymbol{\alpha}_i$(i=1,2,3) represented the mean differences in $\boldsymbol{\eta}$ in the groups 1, 2, and 3 relative to group 4.

6.6   Appendix B

We regressed the variables obtained in the complex EF tasks (WCST and ToL tasks) on the latent variables $\boldsymbol{\eta}$. Again we did this in a 4-group analysis, as we wished to investigate possible age-related changes in these regression equations. Let Y_i denote a given dependent variable, i.e., a WCST or ToL task variable. We specified the following regression equation:

$$Y_i = B_{0i} + B_i * \boldsymbol{\eta}_i + \zeta_i \tag{4}$$

where B_{0i} is the intercept, B_i contains the regression coefficients and ζ_i is a residual (Model A). The mean and variance of Y_i were modeled as

$$\mu_{Yi} = B_{0i} + Bi*\alpha_i \tag{5}$$

$$\sigma^2_{Yi} = B_i \Psi_i B_i^t + \sigma^2_{\zeta I} \tag{6}$$

In this regression model, the variance of the dependent variable (i.e., measures derived from WCST or ToL) is decomposed into a part that is explained by the EF variables, corrected for basic processing speed, and a residual part ($\sigma^2_{\zeta i}$). In this model, we may test the hypothesis that the regression coefficients are equal over the groups, i.e., $B_1 = B_2 = B_3 = B_4$ (Model B). Note that over-group equality of the regression coefficients does not necessarily imply that the amount of explained variance is equal across the groups (specifically, we do not constrain Ψ_i or $\sigma^2_{\zeta i}$ to be equal across the groups). Furthermore, we can investigate whether the regression coefficients are significantly greater than zero (Model C).

Chapter 7 ● Summary and Conclusions

The primary objective of this thesis was to contribute to our understanding of the development of executive function (EF) over the period from age 7 to young-adulthood. EF concerns various cognitive processes that regulate behavior in order to pursue and achieve a future goal. Intact EF is crucial in novel or demanding situations that require rapid and flexible adjustment of behavior to the varying situational demands (Miller & Cohen, 2001). These cognitive processes pertain to the skills that are required to prepare for and to execute behaviors that are necessary in daily life - including planning, organization, strategy use, self-monitoring, and mental representation (e.g., Duncan & Owen, 2000; Stuss, Shallice, Alexander, & Picton, 1995).

Studies of patients with lesions to prefrontal cortex (PFC) suggest that EF depends on the integrity of neural systems involving PFC (e.g., Luria, 1966; Miller & Cohen, 2001). Moreover, based on neurophysiological studies, and on the apparent similarity in task performance between PFC patients and children, it is hypothesized that PFC matures slowly (Dempster, 1992; Diamond, 2002; Stuss, 1992; Welsh, 2002). An increasing number of behavioral studies has demonstrated that EF becomes more efficient as children grow older, and that adult-levels of performance on different EF tasks are attained at different ages during childhood and adolescence (for reviews see Diamond, 2002; Welsh, 2002). These findings are usually interpreted in terms of PFC maturation (e.g., Amso & Casey, 2006; Casey, Tottenham, Liston, & Durston, 2005).

Recent developmental functional MRI studies have advanced our understanding of the biological mechanisms underlying the development of EF. These studies indicated that different regions within PFC underlie different components of goal-directed behavior (e.g., Aron, Robbins, & Poldrack, 2004; Crone, Wendelken, Donohue, & Bunge, 2006; Narayanan et al., 2005). The improvement of EF during childhood and adolescence has been interpreted to reflect gradually less diffuse and more focal PFC activation (e.g., Amso & Casey, 2006; Casey et al., 2005).

In the present thesis, we assessed EF using standard neuropsychological tasks, the Tower of Hanoi (ToH), its distant cousin the Tower of London (ToL), and the Wisconsin Card Sorting Task (WCST). The ToH is a frequently used experimental EF task, which requires the participant to move different-sized disks from a start state to a goal state in the smallest possible number of moves, while abiding by a specified set of rules (Simon, 1975). The ToL is a very similar (but not isomorphic) task (e.g., Welsh & Huizinga, 2001). The ToL requires the participant to move three different-colored balls over three different-sized pegs, in order to reach a pre-specified goal state in the smallest possible number of moves, while abiding by a specified set of rules (Shallice, 1982). Finally, the Wisconsin Card Sorting Task requires the participant to infer one of three different

sorting rules, and the flexible switching between the sorting rules, based on trial-by-trial feedback (Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtis, 1993).

The ToH, the ToL, and the WCST are complex tasks, which require distinct underlying cognitive abilities (e.g., Miyake et al., 2000; Stuss et al., 1995). In a study of healthy young-adults, Miyake et al. (2000) found the three frequently postulated EF components, "Working Memory", "Shifting", and "Inhibition", to be distinct, but moderately correlated, constructs. Moreover, the EF component processes differentially predicted performance on the ToH and the WCST. One of the merits of studying EF components (as distinct entities) is that it facilitates a reasonably precise operationalization of these components in terms of experimental paradigms, which do not involve a context of problem solving, categorization, or planning (Miyake et al., 2000; Stuss et al., 1995). In addition to the examination of complex task performance, the present thesis assessed the development of EF by focusing on different EF components identified by Miyake et al. (2000) (Chapters 3 to 6).

7.7.1   A study on strategy knowledge and learning on the Tower of Hanoi disk-transfer task

In Chapter 2 a study is presented in which a large sample of healthy young-adults performed on an extended version of the ToH (Welsh & Huizinga, 2001). The aim of this study was to examine the effect of administering problems in ascending order of move-length vs. in a random order, and the nature of the participants' strategy knowledge (as indexed by analysis of verbal protocols with regard to four elements of the recursive strategy approach to problem solving). Problem administration did not affect task performance or strategy knowledge. Importantly, individual differences in strategy knowledge contributed to task performance relatively early during the task, such that participants with high scores on strategy knowledge performed better on the ToH. The results suggest that the ability to discover the recursive strategy, needed to successfully solve the ToH, rather than the presence of external cues, contributes to individual differences in task performance. The finding of individual differences in task performance suggests the involvement of different cognitive strategies. To examine the nature of these different strategies, more specific analyses using less complex tasks are needed (see Chapters 3 to 6).

7.7.2   A developmental study of set-switching and set-maintenance on the Wisconsin Card Sorting Task

Previous studies revealed that individual differences in complex task performance can be related to development. That is, children and adults may arrive at the same behavioral

outcomes, but do so using very different cognitive strategies and associated neural pathways (Amso & Casey, 2006). The main goal of Chapter 2 was to examine age-related change of performance on the WCST. Previously, Barceló and Knight (2002) posited that successful WCST performance relies on set-switching and set-maintenance processes. Set-switching refers to the ability to flexibly switch to a new sorting rule in the presence of feedback; set-maintenance refers to the ability to retain the current sorting rule in mind, over varying stimulus conditions, while ignoring irrelevant aspects of the stimuli.

In the present study, WCST performance was scored in terms of set-switching and set-maintenance. Based on previous work, we hypothesized that different EF components underlie set-switching and set-maintenance processes during WCST performance (e.g., Barceló & Knight, 2002; Crone, Ridderinkhof, Worm, Somsen, & Van der Molen, 2004). In Chapter 2 we examined the contribution of three frequently postulated EF component processes, Working Memory, Shifting, and Inhibition (Miyake et al., 2000) to developmental change in WCST performance. Working Memory was interpreted as the set of cognitive processes that temporarily retain information in an accessible state, suitable for carrying out any mental task (Cowan, 1998); Shifting was conceptualized as switching back and forth between multiple tasks (Monsell, 2003); Inhibition was defined as the ability to deliberately inhibit dominant, automatic, or pre-potent responses (Logan & Cowan, 1984). Participants in four age groups (7-year-olds, 11-year-olds, 15-year-olds, and 21-year-olds) were presented with a computerized version of the standard WCST, and three tasks assumed to tap Working Memory, Shifting, and Inhibition. Analysis of WCST performance indicated that with advancing age children become better able to adjust their behavior when the task demands (i.e., the sorting rules) change unexpectedly. This was reflected by different developmental rates of set-switching and set-maintenance abilities: set-switching attained adult level of performance in 11-year-olds; whereas set-maintenance reached adult level of performance in 15-year-olds. In addition, we performed principal component analyses (PCAs) to explore the relation between set-switching and set-maintenance processes during WCST performance. These analyses revealed that in 7-year-olds set-switching and set-maintenance loaded on two separate factors, whereas in 11-, 15-, and 21-year-olds the set-maintenance factor was incorporated in the set-switching factor, resulting in one single factor, which explained most of the variance. The observation of the lack of a set-maintenance factor in the three oldest age groups was interpreted to reflect an increased ability during development to resist the entrance of irrelevant information into working memory. Finally, regression analyses were performed in each age group to understand the relative contribution of Working Memory, Shifting, and Inhibition to the factor structure that emerged from the PCAs. These analyses yielded differential contributions of Working Memory, Shifting and Inhibition to set-switching and set-maintenance processes on the WCST. Specifically, in 7-year-olds, Shifting best predicted the set-switching factor and Inhibition

predicted the set-maintenance factor. In 11-year-olds, Shifting best predicted the single factor found in the three oldest age groups, whereas in 15-year-olds Shifting and Working Memory best predicted this factor, and in 21-year-olds this was the case for Working Memory. The results suggested different developmental rates of Working Memory, Shifting and Inhibition, which taken together underlie the increased ability during development to keep different sorting rules in mind, while searching for a new sorting rule.

7.7.3   Two developmental studies on cognitive flexibility

The task-switching paradigm provides a relatively pure measure for assessing the cognitive process of flexibly switching between multiple tasks. The task-switching paradigm requires the participant to switch between two simple tasks (e.g., deciding whether the color of the stimulus is red vs. blue (task A), or deciding whether the shape of the stimulus is a circle vs. a square (task B)). The presentation of trials in mixed blocks allows for a comparison of performance on task repetitions (i.e., A*A* or B*B*) and task alternations (i.e., A*B* or B*A*). Compared to repetition trials, alternation trials responses are typically observed to be slower and less accurate (i.e., switch costs; Monsell, 2003; Rubinstein, Meyer, & Evans, 2001).

Chapters 4 and 5 address the age-related change in task-switching abilities. The main goal of Chapter 4 was to contrast the developmental trends in the ability to flexibly switch between choice tasks, and the ability to flexibly switch from response inhibition (i.e., stopping) to response execution (i.e., going). Participants in three age groups (7-year-olds, 11-year-olds, and 21-year olds) performed on a hybrid task that required them to consider the color (blue or yellow) or shape (triangle or circle) of a target stimulus, and to execute a choice reaction on some trials, but a disjunctive reaction on others. We observed two different developmental patterns. First, the costs involved in switching between choice tasks decreased developed into adolescence. Second, the costs in switching from stopping to going decreased during development, and reached adult level already by age 11. These results suggested that the control mechanism, which facilitates flexibly switching between tasks (i.e., enabling the recruitment and execution of a choice reaction), develops relatively slowly. In contrast, the control mechanism involved in flexibly switching from stopping to going (i.e., the ability to release inhibition) develops at a faster rate.

An alternative explanation for the results obtained in Chapter 4 invokes children's greater sensitivity to interference from task-irrelevant stimulus features. That is, the task-switching paradigm usually requires the participants to perform two different tasks on the same stimulus set that (necessarily) consists of two-dimensional stimuli. Thus, a stimulus contains a relevant aspect (i.e., the current task to perform), and an irrelevant aspect (i.e.,

the task to ignore). For example, target stimuli can either be congruent (i.e., the relevant and the irrelevant level require the same overt response) or incongruent (i.e., the relevant and the irrelevant level require a different overt response). Previous research has suggested that children have greater difficulty resisting interference of irrelevant information (e.g., Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Ridderinkhof, Van der Molen, Band, & Bashore, 1997). This leads to the question whether the observed deficit in task-switching abilities in children is related to greater sensitivity to interference of task-irrelevant stimulus features.

To address this question, we examined the development of the ability to switch from global to local, and from local to global levels of hierarchical stimuli. In Experiment I, reported in Chapter 5, participants in three groups of participants (7-year-olds, 11-year-olds, and 21-year olds) performed on a task that required them to analyze the target stimulus either at a global level or at a local level. The target stimuli consisted of large (global) rectangles/squares that were comprised of smaller (local) rectangles or squares. The use of hierarchical stimuli also allowed for the examination of the effect of interference. Geometric cues consisting of shapes of the exact shape and same size as the target stimuli indicated which task to perform. The results indicated a global precedence effect in all age groups, but this effect did not discriminate between age groups. Interestingly, no trend towards local processing during development was observed. Moreover, in all age groups, a higher cost of switching levels in incongruent trials than in congruent trials (interference effect) was observed, and these costs were found to be greatest in 7-year-olds.

The main goal of Experiment II was to confirm the results obtained in Experiment I using different cues. In Experiment I, a substantial global precedence effect was found in all age groups. Previous research has shown that the strength of cues can affect performance. We suggested that the observed pronounced global precedence effect in Experiment I was related to the relatively easy-to-match geometric cues. Therefore, in Experiment II, we replaced the geometric cues by symbolic cues, aiming at the actual mental processing of the global and local levels of the stimuli, rather than spatially matching the target stimuli to the cues. The cues in the Experiment II were elephants to indicate the "respond-to-global" task, and mice to indicate the "respond-to-local" task. Different participant in the same three age groups (7-year-olds, 11-year-olds, and 21-year olds) performed on the same task as used in Experiment I. Again, the results yielded a global precedence effect in all age groups, but no trend towards local processing was observed. In addition, switching between target levels was associated with cost in all age groups, but this cost was greatest in 7-year-olds. Moreover, we observed a greater sensitivity in 7-year-olds to interference of the irrelevant target level, but this did not further affect the ability to flexibly switch between target levels. The results obtained in Chapter 5 suggested that the ability to flexibly switch between tasks attains young-adult

level in 11-year-olds, and that the features of the presented cue modulate this effect, as indicated by an interaction with the sensitivity to interference, when the cue and target are easy to match.

7.7.4   A developmental study of EF component processes Working Memory, Shifting, and Inhibition, and their contribution to complex EF tasks

We learned from the results presented in Chapters 2 to 5 that individual differences in complex EF task performance can be explained largely in terms of the differential development of underlying EF components. This now leads to the question how EF is organized in children, and whether this organization changes across development. A major theoretical issue related to this question involves the construct validity of EF. Recently, Miyake et al. (2000) used structural equation modeling (SEM) to examine the organization of EF. SEM involves the extraction of variance that is common to multiple tasks to measure EF (i.e., a latent variables approach). Based on their results, Miyake et al. (2000) reported EF to be a multi-dimensional construct (i.e., involving discrete sub-components Working Memory, Shifting, and Inhibition), rather than a one-dimensional construct. One important advantage of the latent variable approach concerns the reduction of the task impurity problem, which refers to the difficulty in interpretation observed test scores, because of contamination by random error and systematic error. A latent variables approach is especially informative in developmental studies. The central objective in Chapter 6 was to examine the developmental trajectories of the EF components Working Memory, Shifting, and Inhibition, and their relation to performance on the WCST and the ToL. Participants in four large age groups (7-, 11-, 15-, and 21-year-olds) were presented with nine basic experimental tasks (three tasks for each EF component), the WCST, and the ToL. First, standard analyses of (co)variance were performed to examine developmental trends in individual EF tasks, while correcting for basic processing speed. Second, confirmatory factor analysis was performed to examine i) the organization of EF in children and young-adults by investigating whether the indicators of the Working Memory, Shifting, and Inhibition tasks measured the same constructs across age; ii) whether this organization changes across development; and iii) how EF component processes contribute to WCST and the ToL the performance across age groups. The study reported in Chapter 6 yielded several results, after correction for within-group individual differences in basic processing speed. First, at a manifest level, analyses of (co)variance revealed the development of WCST and ToL performance, and of the EF components into adolescence. Second, confirmatory factor analysis yielded two common factors: Working Memory and Shifting. However, the variables assumed to tap Inhibition proved to be unrelated. The organization of EF components was observed to remain stable during

development. Third, at a latent level, the development of Shifting continued into adolescence, whereas Working Memory developed into young-adulthood. Finally, regression analyses revealed that the latent Working Memory factor contributed most strongly to WCST performance in all age groups, whereas ToL performance was predicted by one of the tasks assumed to tap Inhibition (i.e., the Stroop task), but only in 21-year-olds. Chapter 6 presents a normative and differentiated picture of EF development from childhood through adulthood. We found support for non-unitary nature of EF. We clearly established the different developmental rates of the EF components Working Memory and Shifting. In all age groups, the tasks used to tap Inhibition failed to load on a latent factor. The current findings correspond well with recent results of brain imaging studies and neurophysiological studies, which indicate differential and relatively slow development of the PFC areas that presumably underlie the EF components examined in this study (e.g., Amso & Casey, 2006; Casey et al., 2005).

7.7.5 Conclusions

Analysis of standard neuropsychological EF tasks suggested that individual differences in performance can be attributed to different EF component processes (Chapters 2 and 3). Therefore, in this thesis the developmental trajectory of EF was examined using more specific analyses, focusing on the more specific EF components Working Memory, Shifting, and Inhibition (Chapters 4 to 6).

First, this thesis added a developmental dimension to the work of Miyake et al. (2000). In a sample of college volunteers Miyake et al. (2000) found the EF components Working Memory, Shifting, and Inhibition to be clearly separable and distinct. In the present thesis, the EF components Working Memory and Shifting were found to be separable at different ages from childhood to young-adulthood (i.e., at ages 7, 11, 15, and 21), as indicated by two latent factors that represented the respective EF components. In contrast to Miyake et al. (2000), a latent Inhibition factor was not observed during this developmental period. Importantly, the organization of these EF component processes remained constant during from childhood to young-adulthood.

Second, the present thesis provided further evidence for the differential development of EF components from childhood to young-adulthood. The developmental trajectories for Working Memory, Shifting, and Inhibition observed in this thesis were consistent with several developmental studies that indicated distinct developmental trends of different EF component processes. Shifting and performance on two Inhibition tasks (i.e., the Stop-signal and Eriksen Flankers tasks) reached mature levels by adolescence, while Working Memory continued to develop into young-adulthood. Importantly, the reported developmental trends cannot be explained in terms of a global change in basic processing speed.

Third, EF components were found to contribute differentially to standard neuropsychological tasks, namely the WCST and ToL. In 7-, 11-, 15-, and 21-year-olds Working Memory played a significant role in WCST performance. In 21-year-olds, the ability to suppress a pre-potent response was found to be a significant predictor of ToL performance. ToL performance in 7-, 11-, and 15-year-olds was not predicted by the EF components identified in the current thesis.

The present thesis provided a detailed behavioral analysis of the organization and development of EF. The results that emerged from this analysis contribute to theories of PFC maturation by pointing out that EF development co-occurs with the functional maturation of PFC (Dempster, 1992; Stuss, 1992). That is, the observed age-related change in EF performance can be attributed to the protracted maturation of the PFC. Neuroimaging studies have identified anatomically and functionally different subregions of the PFC, presumably underlying the distinct EF components, as currently examined at a behavioral level. Therefore, rather than examining EF development by using standard (complex) neuropsychological tasks involving multi-cognitive processes, a focus on separate EF components, as presented in this thesis, facilitates the interpretation of the observations at a behavioral level in terms of the maturation of anatomically and functionally separate systems within the PFC.

In closing, the present thesis presented a neuropsychological approach to the characterization of the developmental pathway of EF over the period from age 7 through young-adulthood. Experimental versions of standard neuropsychological EF tasks were used, in addition to more specific tasks from the fields of Working Memory, Shifting, and Inhibition. This approach resulted in close examination of the developmental pathway of EF over the period from age 7 through young-adulthood. Neurophysiological research in this age range has indicated a protracted course of maturational processes that enhance neuronal transmission. The results of the experiments presented in this thesis indicated that EF is a multi-faceted construct, and that different EF components mature at different rates during childhood and adolescence. Moreover, this thesis provided insight into the cognitive processes underlying performance on standard neuropsychological tasks.

Future research aimed at increasing the understanding of the nature of EF development, and extending the results of the current neuropsychological approach, would benefit from the integration with knowledge emerging from different areas in the field of developmental cognitive (neuro-) science, such as neuroimaging, computational techniques, and genetics. The results from converging methods will help us understand how anatomically and functionally separate systems within PFC emerge over time, and how this results in improvements in intelligent and social behavior during development.

Summary in Dutch /
Samenvatting in het Nederlands

Het voornaamste doel van dit proefschrift was meer inzicht te verkrijgen in de ontwikkeling van executieve functies (EF) bij kinderen en jongeren in de leeftijd van 7 jaar tot jong-volwassenheid. EF zijn cognitieve processen die het gedrag zodanig reguleren dat het efficiënt en doelgericht kan zijn. EF zijn onontbeerlijk in nieuwe, onbekende situaties waarin een snelle en flexibele aanpassing van gedrag aan de omgeving is vereist (Miller & Cohen, 2001). Deze cognitieve processen zijn nodig in het dagelijks leven, waarin men plannen maakt, organiseert, strategieën hanteert, zichzelf 'in de gaten houdt' en mentale representaties uitvoert (bijv. Duncan & Owen, 2000; Stuss, Shallice, Alexander, & Picton, 1995).

Onderzoek laat zien dat bij patiënten met een beschadiging aan de prefrontale cortex (PFC), het voorste gedeelte van de frontaalkwab van de hersenen, vaak de EF zijn aangetast. Daaruit leidt men af dat EF zijn gerelateerd aan de neurale systemen binnen de PFC (bijv. Luria, 1966; Miller & Cohen, 2001). Ook veronderstelt men dat de PFC zich geleidelijk ontwikkelt. Dat is af te leiden uit de vergelijking tussen neurofysiologische onderzoeken en onderzoek naar EF bij kinderen met onderzoek naar EF in PFC-patiënten, (Dempster, 1992; Diamond, 2002; Stuss, 1992; Welsh, 2002). Een toenemend aantal onderzoeken toont bovendien aan dat EF steeds efficiënter worden naarmate kinderen ouder worden en dat een 'volwassen' niveau van prestatie op verschillende EF-taken wordt bereikt op verschillende momenten tijdens de kindertijd en adolescentie (zie voor overzichtsartikelen Diamond, 2002; Welsh, 2002). Deze bevindingen worden geïnterpreteerd in termen van de ontwikkeling van de PFC (bijv. Amso & Casey, 2006; Casey, Tottenham, Liston, & Durston, 2005).

Dankzij *functional magnetic resonance imaging* (*f*MRI) komen we steeds meer te weten over de biologische mechanismen die ten grondslag liggen aan de ontwikkeling van EF. Uit onderzoek blijkt dat verschillende gebieden in de PFC gerelateerd zijn aan verschillende componenten van efficiënt en doelgericht gedrag (Aron, Robbins, & Poldrack, 2004; Crone, Donohue, Wendelken, & Bunge, 2006; Narayanan et al., 2005). De verbetering van EF tijdens de kindertijd en adolescentie weerspiegelt de activatie van de PFC, waarbij de activiteit als kinderen ouder worden geleidelijk steeds minder diffuus wordt en steeds meer focaal (bijv. Amso & Casey, 2006; Casey et al., 2005).

In dit proefschrift werden EF onderzocht aan de hand van standaard neuropsychologische taken: de *Tower of Hanoi* (ToH), zijn verre neef de *Tower of London* (ToL) en de *Wisconsin Card Sorting Task* (WCST). De ToH is een vaakgebruikte experimentele neuropsychologische taak waarbij de proefpersoon vier schijven van verschillende grootte over drie paaltjes van gelijke hoogte van een begin- naar een

eindstadium moet verplaatsen. Dit moet gebeuren in zo min mogelijk zetten terwijl men zich aan een aantal regels moet houden (Simon, 1975). De ToL lijkt sterk op de ToH maar is desondanks zeer verschillend (Welsh & Huizinga, 2001). Bij de ToL moet de proefpersoon drie verschillend gekleurde bollen verplaatsen over drie paaltjes van verschillende grootte. Ook hier moet een eindstadium worden bereikt in zo min mogelijk zetten terwijl men zich aan een aantal regels houdt (Shallice, 1982). Bij de WCST moet de proefpersoon op basis van feedback per trial drie verschillende sorteerregels ontdekken, gebruiken en afwisselen (Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtis, 1993).

De ToH, de ToL en de WCST zijn echter complexe taken die verschillende onderliggende cognitieve functies vereisen (Miyake et al., 2000; Stuss et al., 1995). In een onderzoek bij gezonde jong-volwassen konden Miyake et al. (2000) drie vaak voorgestelde componenten van EF – werkgeheugen, flexibiliteit en inhibitie – onderscheiden; deze bleken van elkaar afzonderlijk (doch beperkt gecorreleerd). Bovendien werd prestatie op de ToH en de WCST verschillend voorspeld door deze EF-componenten. Een belangrijk voordeel bij het onderzoeken van EF-componenten als afzonderlijke entiteiten ligt in het feit dat deze processen tamelijk nauwkeurig kunnen worden gemeten in experimentele paradigma's die niet vallen binnen een context van probleemoplossen, categoriseren of plannen (Miyake et al., 2000; Stuss et al., 1995). In het onderzoek voor dit proefschrift werd naast de ontwikkeling van prestatie op complexe taken (Hoofdstukken 2 en 3) ook gekeken naar de ontwikkeling van EF, waarbij de aandacht werd gericht op de EF-componenten die door Miyake et al. (2000) werden onderscheiden (Hoofdstukken 3 tot en met 6).

Een onderzoek naar strategiehantering en leren op de *Tower of Hanoi*

Hoofdstuk 2 beschrijft een onderzoek bij een grote groep gezonde jong-volwassenen die een uitgebreide versie van de ToH moesten uitvoeren (Welsh & Huizinga, 2001). Het doel van dit onderzoek was het effect te onderzoeken van de volgorde waarin de items werden aangeboden: de situatie waarin deze in oplopende moeilijkheidsgraad werden gepresenteerd werd vergeleken met de situatie waarin dit gebeurde in willekeurige volgorde. Verder werd in dit onderzoek aan de hand van verbale protocollen gekeken naar de kennis van de recursieve strategie die nodig is voor het uitvoeren van ToH. De volgorde van het aanbieden van de items had geen effect op de taakprestatie en op de kennis van de recursieve strategie. Echter, individuele verschillen in de kennis van de recursieve strategie hadden invloed op de ToH-prestatie relatief vroeg tijdens de taak. Dit was te zien aan het feit dat proefpersonen met een grote kennis van de recursieve strategie beter presteerden op de ToH. De resultaten tonen aan dat individuele verschillen in taakprestatie niet gerelateerd zijn aan de aanwezigheid van externe cues, maar aan de

vaardigheid om de recursieve strategie te ontdekken. De aanwezigheid van individuele verschillen in taakprestatie suggereert de betrokkenheid van verschillende onderliggende cognitieve strategieën. Om de aard van deze strategieën te onderzoeken zijn specifiekere analyses en minder complexe taken noodzakelijk (zie Hoofdstukken 3 tot en met 6).

Een onderzoek naar de ontwikkeling *set–switching*- en *set–maintenance*-vaardigheden op de *Wisconsin Card Sorting Task*

Eerder onderzoek liet zien dat individuele verschillen en de prestatie op complexe taken kunnen worden gerelateerd aan ontwikkeling. Dat wil zeggen dat kinderen en volwassenen hetzelfde gedrag kunnen vertonen, maar dat de onderliggende cognitieve strategieën en de daaraan gerelateerde neurale banen zeer verschillend kunnen zijn (Amso & Casey, 2006). Het voornaamste doel van Hoofdstuk 2 was het onderzoeken van leeftijdsgerelateerde veranderingen in de prestatie op de WCST. In eerder onderzoek stelden Barceló en Knight (2002) dat succesvolle prestatie op de WCST is gerelateerd aan set-switching- en *set-maintenance*-processen. *Set-switching* is de vaardigheid om aan de hand van feedback flexibel te kunnen overstappen naar een nieuwe sorteerregel; *set-maintenance* is de vaardigheid om onder verschillende stimuluscondities de huidige sorteerregel te onthouden en ondertussen de irrelevante aspecten van de stimulus te negeren. Het onderzoek in dit proefschrift beoordeelde WCST-prestatie in termen van *set-switching* en *set-maintenance*. Op grond van eerder onderzoek werd verondersteld dat er verschillende EF-componenten ten grondslag liggen aan *set-switching*- en *set-maintenance*-processen tijdens het uitvoeren van de WCST (bijv. Barceló & Knight, 2002; Crone, Ridderinkhof, Worm, Somsen, & Van der Molen, 2004). Hoofdstuk 2 beschrijft onderzoek naar de bijdrage van de drie EF-componenten werkgeheugen, flexibiliteit en inhibitie (Miyake et al., 2000) aan de leeftijdsgerelateerde verandering in WCST-prestatie. Werkgeheugen is daarbij gedefinieerd als de verzameling van cognitieve processen die het mogelijk maken informatie tijdelijk toegankelijk te houden om zo mentale taken te kunnen uitvoeren (Cowan, 1998); flexibiliteit is het kunnen wisselen tussen verschillende taken (Monsell, 2003); inhibitie is de vaardigheid om dominante, automatische motorische responsen te onderdrukken (Logan & Cowan, 1984). Proefpersonen uit vier verschillende leeftijdsgroepen (7-jarigen, 11-jarigen, 15-jarigen en 21-jarigen) voerden een computerversie van de WCST uit, en drie taken om werkgeheugen, flexibiliteit en inhibitie te meten. Analyse van de prestatie op de WCST liet zien dat kinderen naarmate ze ouder worden, steeds beter in staat zijn hun gedrag aan te passen wanneer de taakvereisten onverwacht veranderen (dat wil zeggen, wanneer de sorteerregels wisselen). Dit was te zien aan de verschillende tempo's waarin *set-switching*- en *set-maintenance*-vaardigheden zich ontwikkelden: een volwassen niveau van *set-switching*-vaardigheden werd op 11-jarige leeftijd bereikt, terwijl een volwassen

niveau van *set-maintenance*-vaardigheden werd bereikt op 15-jarige leeftijd. Verder voerden we principale componenten-analyses (PCAs) uit om de relatie tussen *set-switching* en *set-maintenance* te onderzoeken. Deze analyses lieten zien dat *set-switching* en *set-maintenance* bij 7-jarigen als twee afzonderlijke factoren te onderscheiden zijn. Bij de 11-, 15-, en 21-jarigen was de *set-maintenance*-factor echter niet te onderscheiden van de *set-switching*-factor. Er werd daarom geconcludeerd dat kinderen tijdens hun ontwikkeling steeds beter in staat zijn irrelevante informatie die het werkgeheugen binnenkomt, te negeren.

Ten slotte werden per leeftijdsgroep regressie-analyses uitgevoerd om te onderzoeken wat de relatieve bijdrage was van werkgeheugen, flexibiliteit en inhibitie aan de factorstructuur zoals die naar voren kwam uit de PCA. Deze analyses lieten verschillende patronen zien van de bijdrage van werkgeheugen, flexibiliteit en inhibitie aan *set-switching* en *set-maintenance* tijdens het uitvoeren van de WCST. Bij 7-jarigen werd de *set-switching*-factor het best voorspeld door flexibiliteit en *set-maintenance* door inhibitie. Bij 11-jarigen werd de gevonden gecombineerde factor het best voorspeld door flexibiliteit, bij 15-jarigen door zowel flexibiliteit als werkgeheugen, en bij 21-jarigen door werkgeheugen. De resultaten suggereerden verschillende tempo's in de ontwikkeling van werkgeheugen, flexibiliteit en inhibitie – samen zorgen zij ervoor dat iemand tijdens zijn ontwikkeling steeds beter leert verschillende sorteerregels te onthouden en ondertussen te zoeken naar een nieuwe sorteerregel.

Twee onderzoeken naar de ontwikkeling van cognitieve flexibiliteit

Het *task-switch*-paradigma is een relatief zuivere methode om te onderzoeken hoe flexibel een proefpersoon twee verschillende taken kan afwisselen. Daarbij gaat het om eenvoudige taken, bijvoorbeeld het beslissen of de kleur van een stimulus rood is of blauw (taak A), of het beslissen of de vorm van een stimulus een cirkel of een vierkant is (taak B). Het aanbieden van taken in gemixte blokken maakt het mogelijk de prestatie op taak-repetities (ie., A*A* of B*B*) te vergelijken met de prestatie op taak-alternaties (ie., A*B* of B*A*). Prestatie op taak-alternaties is ten opzichte van taak-repetities altijd langzamer en minder accuraat. Er zijn dus kosten verbonden aan taak-switchen (Monsell, 2003; Rubinstein, Meyer, & Evans, 2001).

Deze vaardigheid in het flexibel wisselen tussen taken komt aan de orde in Hoofdstukken 4 en 5. Hoofdstuk 4 behandelt de ontwikkeling van het vermogen tot wisselen tussen keuzetaken en en tot wisselen van respons-inhibitie (i.e., 'stoppen') en respons-executie (i.e., 'gaan'). Proefpersonen uit drie leeftijdsgroepen (7-jarigen, 11-jarigen en 21-jarigen) voerden een hybride taak uit die het maken van onderscheid tussen kleuren (blauw of geel) en tussen vormen (driehoek of cirkel) vereiste. Sommige trials vroegen om een keuze-reactie en andere een disjunctieve reactie waarbij één van twee

keuze-reacties moest worden gestopt. Daarbij kwamen twee verschillende ontwikkelingspatronen naar voren. Ten eerste namen de kosten van het wisselen tussen keuze-reacties tijdens de ontwikkeling af. Daarbij werd een volwassen niveau van taakprestatie bereikt na de leeftijd van 11 jaar. Ten tweede namen de kosten van het wisselen tussen 'stoppen' en 'gaan' af tijdens de ontwikkeling, waarbij een volwassen niveau van taakprestatie werd bereikt op 11-jarige leeftijd. De resultaten suggereerden dat het controlemechanisme voor het flexibel wisselen tussen taken (i.e., het mechanisme dat het mogelijk maakt een keuze-reactie te recruteren en uit te voeren), zich langzamer ontwikkelt dan het mechanisme dat betrokken is bij het flexibel wisselen van 'stoppen' naar 'gaan' (i.e., het mechanisme dat het mogelijk maakt inhibitie op te heffen).

Een alternatieve verklaring is gerelateerd aan de grotere gevoeligheid van kinderen voor taak-irrelevante onderdelen van de stimulus. Dat wil zeggen, het taak-switch-paradigma vereist dat een proefpersonen twee verschillende taken binnen dezelfde verzameling van stimuli uitvoert. Deze verzameling bestaat (noodzakelijkerwijs) uit twee-dimensionale stimuli. Een stimulus bezit namelijk een relevant aspect (i.e., de taak die moet worden uitgevoerd) en een irrelevant aspect (i.e., de taak die moet worden genegeerd). Een stimulus kan congruent zijn (i.e., het relevante aspect en het irrelevante aspect vereisen dezelfde reactie) of incongruent (i.e., het relevante aspect en het irrelevant aspect vereisen een verschillende reactie). Eerder onderzoek liet zien dat kinderen meer moeite hebben met het ondrukken van irrelevante stimulusinformatie (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Ridderinkhof, Van der Molen, Band, & Bashore, 1997). Dit leidt tot de vraag of het waargenomen verminderde vermogen van kinderen te wisselen tussen taken is gerelateerd aan hun grotere gevoeligheid voor irrelevante stimulusinformatie.

Om deze vraag te beantwoorden werd de ontwikkeling van het vermogen te wisselen tussen het lokale en het globale niveau van hiërarchische stimuli onderzocht. Een hiërarchische stimulus is opgebouwd uit delen (bijv. kleine vierkantjes; het lokale niveau) die samen het geheel vormen (bijv. een groot vierkant; het globale niveau). In Hoofdstuk 5, Experiment I voerden proefpersonen uit drie verschillende leeftijdsgroepen (7-jarigen, 11-jarigen en 21-jarigen) een taak uit waarbij ze de stimulus moesten identificeren op basis van zijn lokale dan wel zijn globale niveau. De stimuli bestonden uit grote (globale) vierkanten/rechthoeken die bestonden uit kleine (lokale) vierkanten/rechthoeken. Het gebruik van deze hiërarchische stimuli maakte het ook mogelijk het effect van interferentie te onderzoeken. Geometrische cues met dezelfde vorm en grootte als de stimuli gaven aan welke taak er moest worden uitgevoerd. De resultaten lieten in alle leeftijdsgroepen een global precedence-effect zien (i.e., snellere en nauwkeuriger reacties op globale' trials), en de grootte van dit effect verschilde niet tussen de groepen. Opvallend was de afwezigheid tijdens de ontwikkeling van een trend richting lokale verwerking. Ten slotte bleken er in alle leeftijdsgroepen hogere kosten op

te treden als er tussen stimulusniveaus moest worden gewisseld, en deze kosten waren hoger op incongruente trials dan op congruente trials, en waren het grootst bij 7-jarigen.

Het belangrijkste doel van Hoofdstuk 5, Experiment II was om de resultaten van Experiment I te bevestigen door het gebruik van symbolische cues. In Experiment I werd in alle leeftijdsgroepen een aanzienlijk global-precedence-effect gevonden. Uit eerder onderzoek blijkt dat de aard van de cues de taakprestatie kan beïnvloeden. In Experiment II werd verondersteld dat het waargenomen global-precendence-effect gerelateerd was aan de relatief makkelijk te matchen geometrische cues. Daarom werden in Experiment II de geometrische cues vervangen door symbolische cues, met als doel het daadwerkelijk mentaal verwerken van de globale en lokale niveaus van de stimuli (in tegenstelling tot het ruimtelijk matchen van de stimuli met de geometrische cues). De cues in Experiment II bestonden uit olifanten om de 'globale' taak aan te duiden, en uit muizen om de 'lokale' taak aan te duiden. Andere proefpersonen uit dezelfde drie leeftijdsgroepen (7-jarigen, 11-jarigen en 21-jarigen) voerden dezelfde taak uit als in Experiment I. Wederom werd in alle leeftijdsgroepen een global-precendence-effect waargenomen, en geen trend richting lokale verwerking. Bovendien bleek het wisselen tussen taken geassocieerd te zijn met switchkosten, en deze waren het grootst bij 7-jarigen. Ten slotte werd bij 7-jarigen een grotere gevoeligheid voor interferentie van het irrelevante stimulusniveau gevonden, maar dit effect had geen invloed op de vaardigheid om flexibel te kunnen switchen tussen taken. De resultaten van Hoofdstuk 5 suggereren dat een volwassen niveau van het vermogen flexibel te wisselen tussen taken wordt bereikt op 11-jarige leeftijd. Bovendien beïnvloeden de kenmerken van de cues deze bevinding, zoals werd bevestigd door een interactie met de gevoeligheid voor interferentie als de de cue en de stimulus eenvoudig te matchen zijn.

Een onderzoek naar de ontwikkeling van de EF–componenten werkgeheugen, flexibiliteit en inhibitie, en hun bijdrage aan de prestatie op complexe EF–taken

De resultaten van de Hoofdstukken 2 tot en met 5 lieten zien dat individuele verschillen in de prestatie op complexe EF-taken voor een aanzienlijk deel kunnen worden verklaard in termen van de verschillen in tempo's van ontwikkeling van de onderliggende EF-componenten. Dit leidt tot de vraag hoe EF bij kinderen is georganiseerd en of deze organisatie verandert tijdens de ontwikkeling. Een belangrijk theoretisch punt hierbij is de constructvaliditeit van EF. Recentelijk gebruikten Miyake et al. (2000) *structural equation modeling* (SEM) om de organisatie van EF te onderzoeken. SEM extraheert de variantie die wordt gedeeld door verschillende EF-taken (i.e., latente-factorenmethode). Miyake et al. (2000) lieten zien dat EF een multi-dimensioneel construct is (i.e., de afzonderlijke EF-componenten werkgeheugen, flexibiliteit en inhibitie zijn op latent

niveau van elkaar te onderscheiden). Een belangrijk voordeel van de latente-factorenmethode is dat het probleem van taakonzuiverheid kan worden gereduceerd. Met taakonzuiverheid wordt bedoeld dat de interpretatie van geobserveerde testscores wordt bemoeilijkt door het feit dat deze vervuild zijn met random en systematische meetfouten. De latente-factorenmethode is mede hierdoor zeer geschikt voor ontwikkelingsonderzoek. Het hoofddoel van Hoofdstuk 6 was te onderzoeken hoe de ontwikkeling van werkgeheugen, flexibiliteit en inhibitie verloopt en hoe hun relatie is tot prestatie op de WCST en de ToL. Proefpersonen uit vier grote leeftijdsgroepen (7-jarigen, 11-jarigen, 15-jarigen en 21-jarigen) voerden negen experimentele taken uit (drie voor elke EF-component), alsmede de WCST en de ToL. Ten eerste werden standaard (co-)variantie-analyses uitgevoerd om de ontwikkelingsverlopen binnen elke taak te onderzoeken – met correctie voor de algemene verwerkingssnelheid. Ten tweede werden confirmatieve factoranalyses uitgevoerd om: i) de organisatie van EF bij kinderen en jong-volwassenen te onderzoeken door te kijken of de werkgeheugen-, flexibiliteits- en inhibitie-taken dezelfde constructen maten binnen elke leeftijdsgroep; ii) te onderzoeken of deze organsatie verandert tijdens de ontwikkeling; en iii) te onderzoeken hoe EF-componenten bijdragen aan prestatie op de WCST en de ToL tijdens de ontwikkeling.

Het onderzoek in Hoofdstuk 6 leverde een aantal resultaten op. Ten eerste lieten (co-)variantie-analyses op manifest niveau zien dat de prestatie op de WCST, ToL, en de EF-taken tot in de adolescentie ontwikkelt. Ten tweede lieten de confirmatieve factor-analyses twee latente factoren zien: werkgeheugen en flexibiliteit. De variabelen die bedoeld waren om inhibitie te meten bleken ongerelateerd. Verder was de organisatie van EF stabiel tijdens de ontwikkeling. Ten derde bleek de latente factor flexibiliteit zich te ontwikkelen tot in de adolescentie en de latente factor werkgeheugen tot in de jong-volwassenheid. Ten slotte lieten regressie-analyses zien dat de latente factor werkgeheugen in alle leeftijdsgroepen het sterkst bijdroeg aan prestatie op de WCST, terwijl de prestatie op de ToL het best voorspeld werd door de prestatie op één van de inhibitietaken (i.e., de Stroop-taak) – echter alleen bij de 21-jarigen.

Hoofdstuk 6 geeft een normatief en gedifferentieerd beeld weer van de ontwikkeling van EF van de kindertijd tot de jong-volwassenheid. Er werd bewijs gevonden voor de gedifferentieerde aard van EF. Bovendien konden de verschillende tempo's van de ontwikkeling van werkgeheugen en flexibiliteit duidelijk worden vastgesteld. In alle leeftijdsgroepen werd geen latente inhbitiefactor gevonden. De huidige resultaten komen goed overeen met de resultaten van recente (*f*)MRI- en neurofysiologische onderzoeken, die de differentiële en relatief langzame ontwikkeling van de PFC laten zien. Deze ligt hoogstwaarschijnlijk ten grondslag aan de EF-componenten die in het huidige onderzoek werden onderzocht (bijv. Amso & Casey, 2006; Casey et al., 2005).

Conclusies

De analyse van standaard neuropsychologische tests gaf aan dat individuele verschillen in taakprestatie kan worden toegeschreven aan verschillende EF-componenten (Hoofdstukken 2 en 3). In dit proefschrift werd het ontwikkelingsverloop van EF onderzocht, waarbij gebruik werd gemaakt van specifiekere analyses gericht op de specifieke componenten werkgeheugen, flexibiliteit en inhibitie (Hoofdstukken 4 tot en met 6).

Ten eerste werd het onderzoek van Miyake et al. (2000) in een ontwikkelingsperspectief gezet. Miyake et al. (2000) testte een groep gezonde jong-volwassen en liet zien dat de EF-componenten werkgeheugen, flexibiliteit en inhibitie duidelijk afzonderlijke constructen zijn. Het onderzoek in dit proefschrift toont aan dat de (latente) EF-componenten werkgeheugen en flexibiliteit op verschillende momenten tussen de kindertijd en jong-volwassenheid (i.e., op 7-jarige, 11-jarige, 15-jarige, en 21-jarige leeftijd) van elkaar te onderscheiden zijn. In tegenstelling tot Miyake et al. (2000) werd er geen latente inhibitie-factor gevonden. Van belang is tevens dat de organisatie van de EF-componenten stabiel bleef tijdens de ontwikkeling van de kindertijd tot de jong-volwassenheid.

Ten tweede leverde het onderzoek in dit proefschrift aanvullend bewijs voor de verschillende tempo's waarin EF-componenten zich ontwikkelen. De ontwikkelingsverlopen van werkgeheugen, flexibiliteit en inhibitie kwamen overeen met verscheidene ontwikkelingsstudies die verschillende ontwikkelingsverlopen van EF-componenten aantoonden. Flexibiliteit en de prestatie op twee inhibitietaken (i.e., de Stop-signal- en de Eriksen Flankers-taken) bereikten jong-volwassen niveau tijdens de adolescentie, terwijl de ontwikkeling van werkgeheugen zich voortzette tot in de jong-volwassenheid. Van groot belang hierbij is dat deze ontwikkelingsverlopen niet kunnen worden verklaard in termen van een globale verandering van de algemene verwerkingssnelheid.

Ten derde toonde het onderzoek aan dat de bijdrage van EF-componenten aan de prestatie op standaard neuropsychologische taken (de WCST en de ToL) verschilde per standaard taak. Bij 7-, 11-, 15- en 21-jarigen speelde werkgeheugen de belangrijkste rol bij de prestatie op de WCST. Bij 21-jarigen speelde het vermogen een motorische respons te onderdrukken de grootste rol in de prestatie op de ToL. De prestatie op de ToL bij 7-, 11-, 15-jarigen werd niet voorspeld door één van de EF-componenten die in dit proefschrift werden onderzocht.

Dit proefschrift levert een gedetailleerde analyse van de organisatie en ontwikkeling van EF. De resultaten dragen bij aan theorieën over de ontwikkeling van de PFC door aan te geven dat de ontwikkeling van EF samengaat met de functionele ontwikkeling van de PFC (Dempster, 1992; Stuss, 1992). De huidige waargenomen

leeftijdsgerelateerde veranderingen in EF kunnen worden toegeschreven aan de relatief langzame ontwikkeling van de PFC. MRI-onderzoek heeft aangetoond dat de PFC uit anatomisch en functioneel van elkaar te onderscheiden subregio's bestaat, die hoogstwaarschijnlijk ten grondslag liggen aan de EF-componenten die in dit proefschrift op gedragsniveau werden onderzocht. Daarom maakt de methode die hier werd gepresenteerd (i.e., een focus op van elkaar te onderscheiden EF-componenten in plaats van op standaard neuropsychologische tests) het mogelijk om de gevonden resultaten te interpreteren in termen van anatomisch en functioneel onderscheidbare systemen in de PFC.

Ter afsluiting, dit proefschrift liet een neuropsychologische benadering zien van onderzoek naar het ontwikkelingsverloop van EF tussen 7-jarige leeftijd en jong-volwassenheid. Hiervoor werden experimentele versies van standaard neuropsychologische EF-taken gebruikt, evenals specifiekere taken om werkgeheugen, flexibiliteit en inhibitie te meten. Deze benadering maakte een gedetailleerd onderzoek naar de ontwikkeling van EF tussen de kindertijd en jong-volwassenheid mogelijk. Neurofysiologisch onderzoek bij proefpersonen op deze leeftijden lieten een relatief lange ontwikkeling zien van processen die neurale transmissie bevorderen. De resultaten van dit proefschrift laten zien dat EF een construct is dat bestaat uit meerdere componenten, en dat deze componenten zich ontwikkelen in verschillende tempo's tijdens de kindertijd en adolescentie. Verder verschafte dit proefschrift meer inzicht in de cognitieve processen die ten grondslag liggen aan standaard neuropsychologische taken.

Toekomstig onderzoek dat zich richt op de ontwikkeling van EF, en dat daarmee aansluit op de resultaten van de huidige neuropsychologische benadering, kan baat hebben bij de integratie van kennis die voorkomt uit verschillende gebieden van de cognitieve (neuro-)wetenschappen zoals MRI, computationele technieken en genetica. De resultaten van convergerende onderzoeksmethoden zullen ons het ontwikkelingsverloop van anatomisch en functioneel van elkaar te onderscheiden gebieden binnen de PFC helpen begrijpen – en hoe dit bijdraagt aan de ontwikkeling van intelligent en sociaal gedrag.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Altmann, E.M., & Trafton, J.G. (2002). Memory for goals: an activation-based model. *Cognitive Science, 26,* 39-83.

Amso, D., & Casey, B. J. (2006). Beyond what develops when - Neuroimaging may inform how cognition changes with development. *Current Directions in Psychological Science, 15*(1), 24-29.

Anderson, J.R., & Douglass, S. (2001). Tower of Hanoi: Evidence of the cost of goal retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1331-1346.

Anderson, S. W., Damasio, H., Jones, R. D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology, 13*(6), 909-922.

Anderson, K., Byrd, D., & Berg, W. K. (2005). The development of the contingent negative variation during the Tower of London: Planning from childhood to early adulthood. *Journal of Cognitive Neuroscience*, 59-59.

Anderson, P., Anderson, V., & Lajoie, G. (1996). The Tower of London Test: Validation and standardization for pediatric populations. *Clinical Neuropsychologist, 10*(1), 54-65.

Anderson, V.A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology, 20*(1), 385-406.

Anderson, S.W., Damasio, H., Jones, R. D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology, 13*(6), 909-922.

Andres, P., & Van der Linden, M. (2001). Supervisory attentional system in patients with focal frontal lesions. *Journal of Clinical and Experimental Neuropsychology, 23*(2), 225-239.

APA: *Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR).* Washington DC, American Psychiatric Association; 2000.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience, 6*(2), 115-116.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8*(4), 170-177.

Baddeley, A. (1986). *Working memory*. New York, NY: Oxford University Press.

Bagley, A., Welsh, M. C., Retzlaff, P., Wolf, C., & Bryan, E. (2002). Towers of Hanoi and London: Contribution of procedural learning and inhibition. *Journal of the International Neuropsychological Society, 8*(2), 229.

Baker, K., Segalowitz, S. J., & Ferlisi, M. C. (2001). The effect of differing scoring methods for the Tower of London task on developmental patterns of performance. *Clinical Neuropsychologist, 15*(3), 309-313.

Band, G. P. H., Van der Molen, M. W., Overtoom, C. C. E., & Verbaten, M. N. (2000). The ability to activate and inhibit speeded responses: Separate developmental trends. *Journal of Experimental Child Psychology, 75*(4), 263-290.

Barceló, F. (1999). Electrophysiological evidence of two different types of error in the Wisconsin card sorting test. *Neuroreport, 10,* 1299-1303.

Barceló, F., & Knight, R. T. (2002). Both random and perspective errors underlie WCST deficits in prefrontal patients. *Neuropsychologia, 40*(3), 349-356.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology-General, 132*(1), 71-92.

Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience, 18*(1), 428-437.

Bechara, A., & Martin, E. M. (2004). Impaired decision making related to working memory deficits in individuals with substance addictions. *Neuropsychology, 18*(1), 152-162.

Bédard, A. C., Nichols, S., Barbosa, J. A., Schachar, R., Logan, G. D., & Tannock, R. (2002). The development of selective inhibitory control across the life span. *Developmental Neuropsychology, 21*(1), 93-111.

Berg, E. A. (1948). A simple objective test for measuring flexibility in thinking. *Journal of General Psychology, 39,* 15-22.

Best, J.B. (1990). Knowledge acquisition and strategic action in "Mastermind" problems. *Memory and Cognition, 18*(1), 54-64.

Beveridge, M., Jarrold, C., & Pettit, E. (2002). An experimental approach to executive fingerprinting in young children. *Infant and Child Development, 11*(2), 107-123.

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*(1), 48-71.

Blair, C. (2003). Behavioral inhibition and behavioral activation in young children: Relations with self-regulation and adaptation to preschool in children attending head start. *Developmental Psychobiology, 42*(3), 301-311.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53*(1), 605-634.

Borys, S. V., Spitz, H. H., & Dorans, B. A. (1982). Tower of Hanoi performance of retarded young adults and nonretarded children as a function of solution length and goal state. *Journal of Experimental Child Psychology, 33*(1), 87-110.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26*(2), 571-593.

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45*(4), 743-754.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Development of frontal lobe contributions to cognitive control in children: Evidence from *f*MRI. *Neuron, 33*, 301-311.

Burack, J. A., Enns, J. T., Iarocci, G., & Randolph, B. (2000). Age differences in visual search for compound patterns: Long- versus short-range grouping. *Developmental Psychology, 36*(6), 731-740.

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society, 4*(6), 547-558.

Carlin, D., Bonerba, J., Phipps, M., Alexander, G., Shapiro, M., & Grafman, J. (2000). Planning impairments in frontal lobe dementia and frontal lobe lesion patients. *Neuropsychologia, 38*(5), 655-665.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology, 40*(6), 1105-1122.

Case, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.

Case, R., Kurland, D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*(3), 386-404.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54*, 241-257.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences, 9*(3), 104-110.

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., et al. (1997). A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. *Journal of Cognitive Neuroscience, 9*(6), 835-847.

Caviness, V. S., Kennedy, D. N., Richelme, C., Rademacher, J., & Filipek, P. A. (1996). The human brain age 7-11 years: A volumetric analysis based on magnetic resonance images. *Cerebral Cortex, 6*(5), 726-736.

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental-Psychology, 37*(5), 715-730.
Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life-span. *Acta Psychologica, 86*(2-3), 109-197.

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology, 8*(3), 219-228.

Chelune, G. J., & Thompson, L. L. (1987). Evaluation of the general sensitivity of the Wisconsin Card Sorting Test among younger and older children. Developmental Neuropsychology, 3(1), 81-89.

Chugani, H. T., Phelps, M. E., & Mazziotta, J. C. (1987). Positron emission tomography study of human brain functional development. *Annals of Neurology, 22*(4), 487-497.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*(1), 45-77.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 83-103). New York: Guilford.

Cohen, R. L., & Heath, M. (1990). The development of serial short-term-memory and the articulatory loop hypothesis. *Intelligence, 14*(2), 151-171.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547-552.

Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences, 2*(3), 77-78.

Cowan, N., Saults, J. S., & Elliott, E. M. (2002). The search for what is fundamental in the development of working memory. In *Advances in Child Development and Behavior, Vol 29* (Vol. 29, pp. 1-49).

Crone, E. A., Bunge, S. A., Van der Molen, M. W., & Ridderinkhof, K. R. (2006). Switching between tasks and responses: A developmental study. *Developmental Science*, *9* (3), 278-287

Crone, E. A., Donohue, S. E., Wendelken, C., Honomichl, R., & Bunge, S. A. (2006). Prefrontal brain regions contributing to developmental changes in rule use. *Submitted*

Crone, E. A., Ridderinkhof, K. R., Worm, M., Somsen, R. J. M., & Van der Molen, M. W. (2004). Switching between spatial stimulus-response mappings: a developmental study of cognitive flexibility. *Developmental Science, 7*(4), 443-455.

Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A. (2006). Neural evidence for dissociable components of task-switching. *Cerebral Cortex, 16*(4), 475-486.

Culbertson, W. C., & Zillmer, E. A. (1998). The construct validity of the Tower of London(DX) as a measure of the executive functioning of ADHD children. *Assessment, 5*(3), 215-226.

Davis, H. P., & Klebe, K. J. (2001). A longitudinal study of the performance of the elderly and young on the Tower of Hanoi puzzle and Rey recall. *Brain and Cognition, 46*(1-2), 95-99.

De Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Attention and Performance XVIII: Control of Cognitive Performance*. Cambridge, MA: MIT Press.

Delis, D. C., Robertson, L. C., & Efron, R. (1986). Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia, 24*(2), 205-214.

DeLuca, C. R., Wood, S. J., Anderson, V., Bucanan, J., Proffitt, T. M., Mahony, K., et al. (2003). Normative data from the Cantab: I. Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology, 25*(2), 242-254.

Demakis, G. J. (2003). A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology, 17*(2), 255-264.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12*(1), 45-75.

Dempster, F. N. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. In M. L. Howe & R. Pasnak (Eds.), *Emerging themes in cognitive development: Vol. I. Foundations* (pp. 3-27). New York: Springer.

Devine, S., Welsh, M. C., Retzlaff, P., Yoh, M., & Adams, C. (2001). Explicit and implicit cognitive processes underlying Tower of Hanoi performance. *Journal of the International Neuropsychological Society, 7*(2), 250.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466-503). New York: Oxford University Press.

Diamond, A., Briand, L., Fossella, J., & Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *American Journal of Psychiatry, 161*(1), 125-132.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38*(3), 352-362.

Dolan, C. V. (2000). Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis. *Multivariate Behavioral Research, 35*(1), 21-50.

Dolan, C. V., Van der Sluis, S., & Grasman, R. P. P. (2005). A note on normal theory power calculation in SEM with data missing completely at random. *Structural Equation Modeling-a Multidisciplinary Journal, 12*(2), 245-262.

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid Intelligence after frontal-lobe lesions. *Neuropsychologia, 33*(3), 261-268.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*(3), 257-303.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*(10), 475-483.

Durston, S., Thomas, K. M., Yang, Y. H., Ulug, A. M., Zimmerman, R. D., & Casey, B. J. (2002). A neural basis for the development of inhibitory control. *Developmental Science, 5*(4), F9-F16.

Eenshuistra, R. M., Ridderinkhof, K. R., & van der Molen, M. W. (2004). Age-related changes in antisaccade task performance: Inhibitory control or working-memory engagement? *Brain and Cognition, 56*(2), 177-188.

Emick, J., & Welsh, M.C. (2005). Association between formal operational thought and executive function as measured by the Tower of Hanoi-Revised. *Learning and Individual Differences, 15* (3), 177-188.

Enns, J. T., Burack, J. A., Iarocci, G., & Randolph, B. (2000). The orthogenetic principle in the perception of "forests" and "trees"? *Journal of Adult Development, 7*(1), 41-48.

Enns, J. T., & Kingstone, A. (1995). Access to global and local properties in visual-search for compound stimuli. *Psychological Science, 6*(5), 283-291.

References

Eriksen, C. W., & Eriksen, B. A. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics. Perception and Psychophysics, 16, 143–149., 16*, 143-149.

Eslinger, P. J., Flaherty-Craig, C. V., & Benton, A. L. (2004). Developmental outcomes after early prefrontal cortex damage. *Brain and Cognition, 55*(1), 84-103.

Espy, K. A., & Bull, R. (2005). Inhibitory processes in young children and individual variation in short-term memory. *Developmental Neuropsychology, 28*(2), 669-688.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*(3), 340-347.

Filoteo, J. V., Friedrich, F. J., & Stricker, J. L. (2001). Shifting attention to different levels within global-local stimuli: A study of normal participants and a patient with temporal-parietal lobe damage. *Cognitive Neuropsychology, 18*(3), 227-261.

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26*(7), 874-890.

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of. *Human Development, 14*, 272-278.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410-419.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177-190.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action - performance of children 31/2-7 years old on a Stroop-like Day-Night Test. *Cognition, 53*(2), 129-153.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Lui, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience, 2*(10), 861-863.

Glosser, G., & Goodglass, H. (1990). Disorders in executive control functions among aphasic and other brain-damaged patients. *Journal of Clinical and Experimental Neuropsychology, 12*(4), 485-501.

Goel, V., & Grafman, J. (1995). Are the frontal lobes implicated in planning functions - interpreting data from the Tower of Hanoi. *Neuropsychologia, 33*(5), 623-642.

Goel, V., Pullara, S. D., & Grafman, J. (2001). A computational model of frontal lobe dysfunction: Working memory and the Tower of Hanoi task. *Cognitive Science, 25*(2), 287-313.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 101*(21), 8174-8179.

Goldberg, T. E., Saint Cyr, J. A., & Weinberger, D. R. (1990). Assessment of procedural learning and problem solving in schizophrenic patients by Tower of Hanoi type tasks. *Journal of Neuropsychiatry and Clinical Neurosciences, 2*(2), 165-173.

Goldstein, F. C., & Green, R. C. (1995). Assessment of problem solving and executive functions. In R. L. Mapou & J. Spector (Eds.), *Clinical neuropsychological assessment: A cognitive approach. Critical issues in neuropsychology.* (pp. 49 - 81). New York: Plenum Press.

Goto, K., Wills, A. J., & Lea, S. E. G. (2004). Global-feature classification can be acquired more rapidly than local-feature classification in both humans and pigeons. *Animal Cognition, 7*(2), 109-113.

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology, 34*, 404-411.

Hambrick, D. Z., & Engle, R. W. (2003). The role of working memory in problem solving. In J. E. Davidson (Ed.), *(2003). The psychology of problem solving* (pp. 176-206). New York: Cambridge University Press.

Handley, S.J., Capon, A., Copp, & Harper, C. (2002). Conditional reasoning and the Tower of Hanoi: The role of spatial and verbal working memory. *British Journal of Psychology, 93*, 501-518.

Harlow, J. M. (1848). Passage of an iron rod through the head. *Boston medical and surgical journal, 39*, 389-393.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. Brainerd (Eds.), *New perspectives on interference and inhibition in cognition* (pp. 175-204). New York: Academic Press.

Harnishfeger, K. K. (1996). Inhibitory processes in childhood development. *International Journal of Psychology, 31*(3-4), 2665-2665.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). *Wisconsin Card Sorting Test Manual, Revised and Expended*. Odessa: Psychological Assessment Resources.

References

Hill, E. L., & Frith, U. (2003). Understanding autism: insights from mind and brain. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 358*(1430), 281-289.

Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology, 7*(4), 347-362.

Hoffmann, J., Kiesel, A., & Sebald, A. (2003). Task switches under Go/NoGo conditions and the decomposition of switch costs. *European Journal of Cognitive Psychology, 15*(1), 101-128.

Hübner, R. (1997). The effect of spatial frequency on global precedence and hemispheric differences, *Perception & Psychophysics* (Vol. 59, pp. 187-201).

Hübner, R. (2000). Attention shifting between global and local target levels: The persistence of level-repetition effects. *Visual cognition, 7*(4), 465-484.

Huizinga, M., Dolan, C. V., & Van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44, 2017-2036.*

Huizinga, M., & Van der Molen, M. W. (2005). Developmental change in switching from color to shape, and from stopping to going. S*ubmitted*.

Huizinga, M., & Van der Molen, M. W. (2006). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Submitted.*

Huizinga, M., & Van der Molen, M. W. (2006). Developmental changes in random number generation. *In preparation*.

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Research Bulletin, 163*(2), 195-205.

Issacs, E. B., & Vargha-Khadem, F. (1989). Differential course of development of spatial and verbal memory span: A normative study. *British Journal of Developmental Psychology, 7*, 377–380.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 14*(89), 81.

Johnson, M. H. (1998). Developing an attentive brain. In R. Parasuraman (Ed.),   *The Attentive Brain* (pp. 427-444.). Cambridge, MA: MIT Press.

Jonkman, L. M., Lansbergen, M., & Stauder, J. E. A. (2003). Developmental differences in behavioral and event-related brain responses associated with response preparation and inhibition in a go/nogo task. *Psychophysiology, 40*(5), 752-761.

Jöreskog, K. G., & Sörbom, D. (1999). *LISREL 8*. Chicago: Scientific Software International.

Kail, R. (1991). Processing time declines exponentially during childhood and adolescence. *Developmental Psychology, 27*(2), 259-266.

Kemps, E., De Rammelaere, S., & Desmet, T. (2000). The development of working memory: Exploring the complementarity of two models. *Journal of Experimental Child Psychology, 77*(2), 89-109.

Kerr, B., Blanchard, C., & Miller, K. (1980). Children's use of sequence information in partially predictable reaction-time sequences. *Journal of Experimental Child Psychology, 29*(3), 529-549.

Kerr, B., Davidson, J., Nelson, J., & Haley, S. (1982). Stimulus and response contributions to the children's reaction-time repetition effect. *Journal of Experimental Child Psychology, 34*(3), 526-541.

Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Cognitive functions in the prefrontal cortex--Working memory and executive control. *Current Directions in Psychological Science, 6*(6), 185-192.

Kimchi, R. (1988). Selective attention to global and local-levels in the comparison of hierarchical patterns. *Perception & Psychophysics, 43*(2), 189-198.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin, 112*(1), 24-38.

Kimchi, R., Hadad, B., Behrmann, M., & Palmer, S. E. (2005). Microgenesis and ontogenesis of perceptual organization - Evidence from global and local processing of hierarchical patterns. *Psychological Science, 16*(4), 282-290.

Kimchi, R., & Merhav, I. (1991). Hemispheric processing of global form, local form, and texture. *Acta Psychologica, 76*(2), 133-147.

Kimchi, R., & Palmer, S. E. (1985). Separability and integrality of global and local levels of hierarchical patterns, *Journal of Experimental Psychology: Human Perception and Performance* (Vol. 11, pp. 673-688).

Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing: "Top-down," "bottom-up," or "middle-out." *Perception and Psychophysics, 25*(3), 225-231.

Kipp, K. (2005). A developmental perspective on the measurement of cognitive deficits in attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57*(11), 1256-1260.

Kirby, N. H. (1980). Sequential effects in choice reaction time. In A. T. Welford (Ed.), *Reaction times* (pp. 129-127). London: Academic Press.

Kirk, U., & Kelly, M. A. (1986). Children's differential performance on selected dorsolateral prefrontal and posterior cortical functions: a developmental perspective. *Journal of Clinical and Experimental Neuropsychology, 7*, 604.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*(5), 449-467.

References

Kleinsorge, T., & Gajewski, P. D. (2004). Preparation for a forthcoming task is sufficient to produce subsequent shift costs. *Psychonomic Bulletin & Review, 11*(2), 302-306.

Klenberg, L., Korkman, M., & Lahti Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology, 20*(1), 407-428.

Kline, R. T. (1998). *Principles and practice of Structural Equation Modeling*. New York: Guilford.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience, 14*(1), 1-10.

Konishi, S., Chikazoe, J., Jimura, K., Asari, T., & Miyashita, Y. (2005). Neural mechanism in anterior prefrontal cortex for inhibition of prolonged set interference. *Proceedings of the National Academy of Sciences of the United States of America, 102*(35), 12584-12588.

Konishi, S., Kawazu, M., Uchida, I., Kikyo, H., Asakura, I., & Miyashita, Y. (1999). Contribution of working memory to transient activation in human inferior prefrontal cortex during performance of the Wisconsin Card Sorting Test. *Cerebral Cortex, 9*, 745-753.

Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, J., et al. (2005). Development of attentional networks: An fMRI study with children and adults. *Neuroimage, 28*(2), 429-439.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17*(2), 248-294.

Kramer, A. F., Hahn, S., & Gopher, D. (1999). Task coordination and aging: Explorations of executive control processes in the task switching paradigm. *Acta Psychologica, 101*(2-3), 339-378.

Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied, 1*(1), 50-76.

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica, 115*(2-3), 143-165.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15*(1), 126-147

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica, 115*(2-3), 143-165.

Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Academy of Sciences, 99*(20), 13336-13341.

LaGasse, L. L. (1993). Effects of good form and spatial-frequency on global precedence. *Perception & Psychophysics, 53*(1), 89-105.

Lamb, M. R., & Robertson, L. C. (1990). The effect of visual angle on global and local reaction-times depends on the set of visual angles Presented. *Perception & Psychophysics, 47*(5), 489-496.

Lamb, M. R., Robertson, L. C., & Knight, R. T. (1990). Component mechanisms underlying the processing of hierarchically organized patterns - Inferences from patients with unilateral cortical-Lesions. *Journal of Experimental Psychology-Learning Memory and Cognition, 16*(3), 471-483.

Lamb, M. R., & Yund, E. W. (1996). Spatial frequency and interference between global and local levels of structure, *Visual Cognition* (Vol. 3, pp. 193-219).

Lamb, M. R., & Yund, E. W. (2000). The role of spatial frequency in cued shifts of attention between global and local forms. *Perception & Psychophysics, 62*(4), 753-761.

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence, 12*(2), 131-147.

Lehto, J. H. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 49*(1), 29-50.

Lehto, J. H. (2004). A test for children's goal-directed behavior: A pilot study. *Perceptual and Motor Skills, 98*(1), 223-236.

Lehto, J. H., Juujaervi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*(1), 59-80.

Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., et al. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology, 7*(3), 377-395.

Levin, H. S., Fletcher, J. M., Kufera, J. A., Harward, H., Lilly, M. A., Mendelsohn, D., et al. (1996). Dimensions of cognition measured by the Tower of London and other cognitive tasks in head-injured children and adolescents. *Developmental Neuropsychology, 12*(1), 17-34.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). London: Oxford University Press.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

References

Lock, C., Welsh, M. C., Adams, C., & Kurtz, A. (2002). Tower of Hanoi: Influence of strategy instruction and extended practice on performance. *Journal of the International Neuropsychological Society, 8*, 229.

Logan, G. D. (1994). *On the ability to inhibit thought and action: A users' guide to the stop signal paradigm*. San Diego, CA: Academic Press.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance, 29*(3), 575-599.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295-327.

Luce, R. D. (1986). Response times: Their role in inferring elementary mental organization. Oxford: Oxford University Press.

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development, 76*(3), 697-712.

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia, 36*(3), 273-293.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75*(5), 1357-1372.

Luna, B., & Sweeney, J. A. (2004). The emergence of collaborative brain function - fMRI studies of the development of response inhibition. In *Adolescent brain development: Vulnerabilities and opportunities* (Vol. 1021, pp. 296-309).

Luria, A. R. (1966). *Higher cortical function in man*. Oxford, England: Basic Books.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226.

Martin, M. (1979). Local and global processing: The role of sparsity. *Memory and Cognition, 7*(6), 476-484.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255-260.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129*(1), 4-26.

McLean, J. P. (1979). Perspectives on the forest and trees: The precedence of parts and wholes in visual information processing. *Dissertation Abstracts International, 39*, 6162b–6163b.

155

References

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*(6), 1423-1442.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*(3), 211-253.

Meredith, W. (1993). Measurement invariance, factor-analysis and factorial invariance. *Psychometrika, 58*(4), 525-543.

Mesulam, M.-M. (2002). The human frontal lobes: Transcending the default mode through contingent encoding. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 8-30). New York: Oxford University Press.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms. *Psychological Review, 104*(1), 3-65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*(4), 749-791.

Miller, P. H. (1994). Individual differences in childrens strategic behaviors - Utilization deficiencies. *Learning and Individual Differences, 6*, 285-307.

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

Miller, J., & Navon, D. (2002). Global precedence and response activation: Evidence from LRPs. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 55*(1), 289-310.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9*, 100-110.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience, 21*(19), 7733-7741.

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved Mysteries of the Mind: Tutorial Essays in Cognition* (pp. 93–148). New York: Erlbaum.

Monsell, S. (2003). Task-set reconfiguration processes do not imply a control homunculus: Reply to Altmann. *Trends in Cognitive Sciences, 7*(8), 341-342.

156

References

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140.

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research/Psychologische Forschung, 63*(3-4), 250-264.

Morris, R. G., Ahmed, S., Syed, G. M., & Toone, B. K. (1993). Neural correlates of planning ability – frontal lobe activation during the Tower of London test. *Neuropsychologia, 31*(12), 1367.

Muller, U., Zelazo, P. D., Hood, S., Leone, T., & Rohrer, L. (2004). Interference control in a new rule use task: Age-related changes, labeling, and attention. *Child Development, 75*(5), 1594-1609.

Nagahama, Y., Okina, T., Suzuki, N., Nabatame, H., & Matsuda, M. (2005). The cerebral correlates of different types of perseveration in the Wisconsin Card Sorting Test. *Journal of Neurology Neurosurgery and Psychiatry, 76*(2), 169-175.

Narayanan, N. S., Prabhakaran, V., Bunge, S. A., Christoff, K., Fine, E. M., & Gabrieli, J. D. E. (2005). The role of the prefrontal cortex in the maintenance of verbal working memory: An event-related fMRI analysis. *Neuropsychology, 19*(2), 223-232.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353-383.

Navon, D. (1983). How many trees does it take to make a forest? *Perception, 12*(3), 239-254.

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica, 114*(3), 273-309.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical Modeling.* (6th ed.): VCU Box 900126, Richmond, VA 23298: Department of Psychiatry.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220-246.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1-18). New York: Plenum Press.

Numminen, H., Lehto, J. E., & Ruoppila, I. (2001). Tower of Hanoi and working memory in adult persons with intellectual disability. *Research in Developmental Disabilities, 22*(5), 373-387.

Nunally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal-lobe lesions in man. *Neuropsychologia, 28*(10), 1021-1034.

Paniak, C., Miller, H. B., Murphy, D., Patterson, L., & Keizer, J. (1996). Canadian developmental norms for 9 to 14 year-olds on the Wisconsin Card Sorting Test. *Canadian Journal of Rehabilitation, 9*(4), 233-237.

Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. In N. A. Krasnegor & G. R. Lyon (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 265-281). Baltimore: Paul H. Brookes.

Pennington, B. F., Bennetto, L., McAleer, O., & Roberts, R. J., Jr. (1996). Executive functions and working memory: Theoretical and measurement issues. In G. R. Lyon & N. A. Krasnegor (Eds.), *(1996). Attention, memory, and executive function* (pp. 327-348). Baltimore: Paul H. Brookes.

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles, J. M., Zipursky, R. B., & Lim, K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Archives of Neurology, 51*(9), 874-887.

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

Rabbitt, P. M. A., Lowe, C., & Shilling, V. (2001). Frontal tests and models for cognitive ageing. *European Journal of Cognitive Psychology, 13*(1-2), 5-28.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*(3), 510-532.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Oxford: Oxford Psychologists Press.

Raven, J. C., Court, J. H., & Raven, J. (1985). *Raven's Progressive Matrices*. London: J.C. Raven.

Raz, A., & Buhle, J. (2006). Typologies of attentional networks. *Nature Reviews Neuroscience, 7*(5), 367-379.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*(5695), 443-447.

Ridderinkhof, K. R., & Van der Molen, M. W. (1995). When global information and local information collide - a brain potential analysis of the locus of interference effects. *Biological Psychology, 41*(1), 29-53.

Ridderinkhof, K. R., & Van der Molen, M. W. (1995). A psychophysiological analysis of developmental differences in the ability to resist interference. *Child Development, 66*(4), 1040-1056.

Ridderinkhof, K. R., Van der Molen, M. W., Band, G. P. H., & Bashore, T. R. (1997). Sources of interference from irrelevant information: A developmental study. *Journal of Experimental Child Psychology, 65*(3), 315-341.

Rieger, M., & Gauggel, S. (1999). Inhibitory after-effects in the stop signal paradigm. *British Journal of Psychology, 90*(4), 509-518.

Rieger, M., Gauggel, S., & Burmeister, K. (2003). Inhibition of ongoing responses following frontal, nonfrontal, and basal ganglia lesions. *Neuropsychology, 17*(2), 272-282.

Roberts, A.C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex, 10*(3), 252-262.

Roberts, R. J., Jr., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology, 12*(1), 105-126.

Robertson, L. C. (1996). Attentional persistence for features of hierarchical patterns. *Journal of Experimental Psychology-General, 125*(3), 227-249.

Robertson, L. C., Lamb, M. R., & Knight, R. T. (1991). Normal global local analysis in patients with dorsolateral frontal-lobe lesions. *Neuropsychologia, 29*(10), 959-967.

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (Cantab) - a Factor-analytic study of a large-sample of normal elderly volunteers. *Dementia, 5*(5), 266-281.

Rogers, R.D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 763-797.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*(8), 1029-1040.

Rushworth, M., Walton, M., Kennerley, S., & Bannerman, D. (2004). The role of medial frontal cortex in making decisions and changing tasks. *International Journal of Psychology, 39*(5-6), 514-514.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(147-177).

Schmittmann, V. D., Visser, I., & Raijmakers, M. E. J. (2006). Multiple learning modes in the development of performance on a rule-based category-learning task. *Neuropsychologia, 44,*2079-2091.

Schuch, S., & Koch, I. (2003). The role of response selection for inhibition of task sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 92-105.

Schnirman, G. M., Welsh, M. C., & Retzlaff, P. D. (1998). Development of the Tower of London-revised. *Assessment, 5*(4), 355-360.

Scholnick, E. K., Friedman, S. L., & Wallner Allen, K. E. (1997). What do they really measure? A comparative analysis of planning tasks. In S. L. Friedman & E. K.

Scholnick (Eds.), *The developmental psychology of planning: Why, how, and when do we plan?* (pp. 127-156). Mahwah: Lawrence Erlbaum

Schunn, C. D., Lovett, M. C., & Reder, L. M. (2001). Awareness and working memory in strategy adaptivity. *Memory and Cognition, 29*(2), 254-266.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London (biology), 298*, 199-209.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal-lobe damage in man. *Brain, 114*, 727-741.

Shedden, J. M., Marsman, I. A., Paul, M. P., & Nelson, A. (2003). Attention switching between global and local elements: Distractor category and the level repetition effect. *Visual Cognition, 10*(4), 433-470.

Siegel, L. S. (1994). Working memory and reading - a life-span perspective. *International Journal of Behavioral Development, 17*(1), 109-124.

Siegler, R. S. (1983). Information processing approaches to development. In W. Kessen & P. H. Mussen (Eds.), *Handbook of child psychology* (Vol. 1 History, theory, and methods, pp. 129-211). New York: Wiley.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology, 7*, 268-288.

Smith, E. E., & Jonides, J. (1999). Neuroscience - Storage and executive processes in the frontal lobes. *Science, 283*(5408), 1657-1661.

Smulders, S. F. A., Notebaert, W., Meijer, M., Crone, E. A., Van der Molen, M. W., & Soetens, E. (2004). Sequential effects on speeded information processing: A developmental study. *Journal of Experimental Child Psychology, 90 (3)*, 208-234.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174-215.

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 11*(5), 598-616.

Soetens, E., & Hueting, J. (1992). Age influences on sequential effects in serial two-choice reaction time. In H. Bouma & J.A.M. Graafmans (eds.), *Gerontology* (pp. 237-242). Amsterdam. The Netherlands: IOS Press.

Sommer, W., Leuthold, H., & Soetens, E. (1999). Covert signs of expectancy in serial reaction time tasks revealed by event-related potentials. *Perception and Psychophysics, 61*(2), 342-353.

Somsen, R. J. M., Van der Molen, M. W., Jennings, J. R., & Van Beek, B. (2000). Wisconsin Card Sorting in adolescents: Analysis of performance, response times and heart rate. *Acta Psychologica, 104*(2), 227-257.

Sowell, E. R., Delis, D., & Stiles, J. (2001). Improved memory functioning and frontal lobe maturation between childhood and adolescence: a structural MRI study. *Journal of the International Neuropsychological Society, 7*(3), 312-322.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience, 24*(38), 8223-8231.

Span, M. M., Ridderinkhof, K. R., & van der Molen, M. W. (2004). Age-related changes in the efficiency of cognitive processing across the life span. *Acta Psychologica, 117*(2), 155-183.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89*(4), 669-679.

SPSS Inc. (2003). *SPSS Base 12.0 for Windows User's Guide*. SPSS Inc., Chicago IL. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition, 20*(1), 8-23.

Stuss, D. T. (2006). Frontal lobes and attention: Processes and networks, fractionation and integration. *Journal of the International Neuropsychological Society, 12*(2), 261-271.

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin, 95*(1), 3-28.

Stuss, D. T., & Benson, D. F. (1987). The frontal lobes and control of cognition and memory. In E. Perecman (Ed.), *The frontal lobes revisited.* (pp. 141-158). New York: The IRBN Press.

Stuss, D. T., & Levine, B. (2002). Adult clinical neuropsychology: Lessons from studies of the frontal lobes. *Annual Review of Psychology, 53*(1), 401-433.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., et al. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia, 38*(4), 388-402.

Stuss, D. T., & Knight, R. T. (2002). Prefrontal Cortex: The present and future. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 573-597). New York: Oxford University Press.

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. In J. Grafman & K. J. Holyoak (Eds.), *Structure and functions of the human prefrontal cortex. Annals of the New York Academy of Sciences* (pp. 191-211). New York, NY: New York Academy of Sciences.

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 89-103.

Swanson, H. L. (2006). Cognitive processes that underlie mathematical precociousness in young children. *Journal of Experimental Child Psychology, 93*(3), 239-264.

Teuber, H.-L. (1972). Unity and diversity of frontal lobe functions. *Acta Neurobiologiae Experimentalis, 32*, 615-656.

Tranel, D. (2002). Emotion, decision making and the ventromedial cortex. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 338-353). New York: Oxford University Press.

Van Boxtel, G. J. M., Van der Molen, M. W., Jennings, J. R., & Brunia, C. H. M. (2001). A psychophysiological analysis of inhibitory motor control in the stop-signal paradigm. Biological Psychology, 58(3), 229-262.

Van der Molen, M. W. (2000). Developmental changes in inhibitory processing: Evidence from psychophysiological measures. *Biological Psychology, 54*(1-3), 207-239.

Van der Molen, M. W., & Keuss, P. J. (1981). Response selection and the processing of auditory intensity. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 33a*(2), 177-184.

Van der Molen, M. W., & Ridderinkhof, K. R. (1998). The growing and aging brain: Life-span changes in brain and cognitive functioning. In A. Demetriou, W. Doise & C. F. M. Van Lieshout (Eds.), *Life-Span Developmental Psychology: A European Perspective* (pp. 35-100). New York: Wiley.

Van den Wildenberg, W. P. M., & Van der Molen, M. W. (2004). Developmental trends in simple and selective inhibition of compatible and incompatible responses. *Journal of Experimental Child Psychology, 87*(3), 201-220.

Van der Sluis, S., De Jong, P. F., & Van der Leij, A. (2005). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Submitted*

References

Ward, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception and Performance, 8*(4), 562-581.

Welsh, M. C. (1991). Rule-guided behavior and self-monitoring on the Tower of Hanoi disk-transfer task. *Cognitive Development, 6*(1), 59-76.

Welsh, M. C. (2002). Developmental and clinical variations in executive functions. In D. L. Molfese & V. J. Molfese (Eds.), *Developmental variations in learning: Applications to social, executive function, language, and reading skills.* (pp. 139-185). Mahwah: Lawrence Erlbaum.

Welsh, M.C., Cicerello, A., Cuneo, K., & Brennan, M. (1995). Error and temporal patterns in Tower of Hanoi performance: Cognitive mechanisms and individual differences. *Journal of General Psychology, 122*(1), 69-81.

Welsh, M. C., & Huizinga, M. (2001). The development and preliminary validation of the Tower of Hanoi-Revised. *Assessment, 8*(2), 167-176.

Welsh, M. C., Huizinga, M., Granrud, M. A., Cooney, J., Adams, C., & Van der Molen, M. W. (2002). A structural equation model of executive function in normal young adults. *Journal of the International Neuropsychological Society, 8*(2), 264.

Welsh, M. C., & Pennington, B. F. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental Neuropsychology, 4*(3), 199-230.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7*(2), 131-149.

Welsh, M. C., Satterlee Cartmell, T., & Stine, M. (1999). Towers of Hanoi and London: Contribution of working memory and inhibition to performance. *Brain and Cognition, 41*(2), 231-242.

Werner, H. (1948). *Comparative psychology of mental development.* New York: International Universities Press.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development* (pp. 125-148). Minneapolis: University of Minnesota Press.

Wiers, R. W., Gunning, W. B., & Sergeant, J. A. (1998). Is a mild deficit in executive functions in boys related to childhood ADHD or to parental multigenerational alcoholism? *Journal of Abnormal Child Psychology, 26*(6), 415-430.

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology, 35*(1), 205-213.

References

Winter, W. E., Broman, M., Rose, A. L., & Reber, A. S. (2001). The assessment of cognitive procedural learning in amnesia: Why the Tower of Hanoi has fallen down. *Brain and Cognition, 45*(1), 79-96.

Wilson, J. L., Scott, J. H., & Power, K. G. (1987). Developmental differences in the span of visual memory for pattern. *British Journal of Developmental Psychology, 5*(3), 249-255.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology* (Rev. ed.). Oxford, England: Holt.

Wothke, W. (2000). Longitudinal and multi-group modeling with missing data. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wright, I., Waterman, M., Prescott, H., & Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: typical developmental trends. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 44*(4), 561-575.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs." *Psychological Research/Psychologische Forschung, 63*(3-4), 212-233.

Yakovlev, P. I., & Lecours, A. R. (1967). The myelogenetic cycles of regional maturation of the brain. In A. Minkovsky (Ed.), *Regional Development of the Brain in Early Life* (pp. 3-70). Oxford: Blackwell.

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica, 115*(2-3), 167-183.

Zelazo, P. D., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function: Cognitive complexity and control--revised. *Monographs of the Society for Research in Child Development, 68*(3), 93-119.