Assessment and development of executive functions in school-age children

Liisa Klenberg

Institute of Behavioural Sciences
University of Helsinki
Finland

Academic dissertation to be publicly discussed,
by due permission of the Faculty of Behavioural Sciences,
at the University of Helsinki in Auditorium 107, Athena (Siltavuorenpenger 3 A)
on the 20th of March 2015 at 12 o'clock

University of Helsinki
Institute of Behavioural Sciences
Studies in Psychology 108: 2015

Supervisors Professor Marit Korkman, PhD
Institute of Behavioural Sciences
University of Helsinki, Finland

Professor Laura Hokkanen, PhD
Institute of Behavioural Sciences
University of Helsinki, Finland

Reviewers Professor H. Gerry Taylor, PhD
Case Western Reserve University
Rainbow Babies & Children’s Hospital, Cleveland, Ohio, USA

Professor Astri J. Lundervold, PhD
Department of Biological and Medical Psychology
University of Bergen, Norway

Opponent Professor Timo Ahonen, PhD
Department of Psychology
University of Jyväskylä, Finland

ISSN-L 1798-842X
ISSN 1798-842X
ISBN 978-951-51-0866-1 (pbk.)
ISBN 978-951-51-0867-8 (PDF)
http://www.ethesis.helsinki.fi
Unigrafia
Helsinki 2015

Contents

Abstract

Executive functions (EFs) are cognitive processes that direct, coordinate, and control other cognitive functions and behavior. They include processes of inhibition, attention, and self-directed execution of actions. EFs are involved in all purposeful activity, and for children, they are essential for learning and functioning in school environments. Difficulties in EFs are common in school-age children with developmental disabilities, such as attention deficit disorder (ADHD). Understanding normative development forms the necessary basis for the assessment of individual differences in EFs. The developmental findings, however, remain unclear due to methodological challenges in measuring EFs. In the clinical assessment of EF difficulties, EF measures that are sensitive to everyday difficulties are required.

This thesis consists of three studies addressing EFs in school-age children. The first study employed neuropsychological tests from the Developmental Neuropsychological Assessment NEPSY to examine age-related differences in EFs in a sample of 400 children. The second study investigated the methodological factors related to EF measures in a sample of 340 children using response inhibition tasks from the NEPSY-II, the second edition of the Developmental Neuropsychological Assessment. The third study aimed at constructing a new instrument, the Attention and Executive Function Rating Inventory ATTEX, for the clinical assessment of EFs and at verifying the psychometric properties of the rating scale in a sample of 916 children.

Age-related improvement in EF performance continued throughout the school-age period, decelerating at different times in the ten different tasks. The development seemed to proceed from inhibition to attention control, and further to fluency. A closer examination of age-related differences in response inhibition showed developmental variation even within this EF domain. The developmental change in response inhibition was apparent at school age, but the developmental proceeding seemed different when different outcome measures were used. Factors related to the cognitive requirements and the presented stimuli also had an effect on the results.

The ATTEX rating scale demonstrated high internal consistency reliability and good criterion and discriminant validity for ADHD. According to the ATTEX scales, the ADHD subtypes differed from each other in the EF profiles, and children with predominantly inattentive symptoms showed more wide-ranging difficulties in EFs than children with combined symptoms of inattention and hyperactivity-impulsivity.

For both the scientific research and clinical assessment of EFs, carefully examined, reliable, and valid measures are essential. In line with previous studies, this thesis demonstrates that when a large battery of EF tasks is used, the developmental proceeding varies across the different EF components. The relative differences between EF domains, however, may actually reflect the characteristics of the measures more than the EF constructs as such. Numerous factors related to the measures, such as the task materials and stimuli, the outcome measures, the involvement of other cognitive processes, and task sensitivity to age-related difference, have critical effects on the developmental results. For the clinical assessment of EF difficulties, the newly constructed EF rating scale proved to be a suitable measure both for screening and examining the detailed EF profiles of children in school situations.

Tiivistelmä

Toiminnanohjauksella tarkoitetaan kognitiivisia prosesseja, jotka suuntaavat, yhdistävät ja kontrolloivat muita kognitiivisia toimintoja ja käyttäytymistä. Niihin luetaan joukko inhibition, tarkkaavuuden ja itseohjautuvan toiminnan prosesseja. Toiminnanohjausta tarvitaan kaikessa tavoitteellisessa toiminnassa ja, etenkin lapsilla, oppimisessa sekä kouluympäristössä toimimisessa. Toiminnanohjauksen vaikeuksia esiintyy yleisesti lapsilla, joilla on kehitykseen liittyviä häiriöitä kuten tarkkaavuushäiriöitä. Toiminnanohjauksen normaalin kehityksen tunteminen luo välttämättömän perustan vaikeuksien ja yksilöllisen vaihtelun arvioimista varten. Kehitystutkimusten tulokset ovat kuitenkin toistaiseksi jääneet hajanaisiksi tutkimusmenetelmiin liittyvien haasteiden vuoksi. Toiminnanohjauksen vaikeuksin kliinisen arvioinnin haasteena on arjen taitoja luotettavasti kuvaavien arviointimenetelmien vähäisyys.

Tämä väitöstutkimus käsittää kolme kouluikäisten lasten toiminnanohjausta koskevaa tutkimusta. Ensimmäisessä tutkimuksessa toiminnanohjauksen kehitystä arvioitiin 400 lapsen aineistossa Lasten neuropsykologinen tutkimus NEPSY:n osatestejä käyttäen. Toisessa tutkimuksessa selvitettiin arviointimenetelmien metodologisten ominaisuuksien vaikutuksia käyttäen uuden testiversion, NEPSY-II:n inhibitiotehtäviä 340 lapsen kehityksen arvioinnissa. Kolmannen tutkimuksen tavoitteena oli kehittää uusi kyselymenetelmä, Keskittymiskysely, lasten toiminnanohjauksen vaikeuksien kliinistä arviointia varten sekä arvioida menetelmän psykometrisia ominaisuuksia 916 lapsen aineistossa.

Toiminnanohjauksessa havaittiin kehitystä koko kouluiän ajan siten, että kehityksen hidastuminen vaihteli arvioitujen kymmenen testisuorituksen välillä. Kehitys näytti etenevän inhibition kautta tarkkaavuuteen ja lopulta sujuvaan tuottamiseen. Reaktioinhibition kehityksen tarkempi tarkastelu osoitti kehityksen vaihtelevan myös tämän taidon sisällä. Kehityksellisiä muutoksia ilmeni kouluiässä edelleen, mutta kehityksen eteneminen oli eri tulosmuuttujilla mitattuna erilaista. Tehtävien kognitiiviset vaatimukset ja käytetyt ärsykkeet vaikuttivat niin ikään kehityskulkua kuvaaviin tuloksiin.

Keskittymiskyselyn sisäinen reliabiliteetti, kriteerivaliditeetti sekä erottelukyky tarkkaavuushäiriön diagnosoinnissa osoittautuivat vahvoiksi. Kyselyn perusteella tarkkaavuushäiriön alatyyppien toiminnanohjauksen profiilit erosivat toisistaan. Pääsääntöisesti tarkkaamattomilla lapsilla oli laaja-alaisempia toiminnanohjauksen vaikeuksia kuin lapsilla, joilla oli sekä ylivilkkauteen ja impulsiivisuuteen että tarkkaamattomuuteen liittyviä oireita.

Huolellisesti tutkitut, luotettavat ja validit arviointimenetelmät ovat välttämättömiä niin tieteellisessä tutkimuksessa kuin kliinisessä työssäkin. Aikaisempien tutkimusten tavoin tämä väitöstutkimus osoitti, kuinka kehityksen eteneminen eri toiminnanohjauksen toiminnoissa vaihtelee, kun arvioinnissa käytetään laajaa testijoukkoa. Eri toimintojen välillä havaitut kehityserot voivat kuitenkin pikemminkin heijastaa menetelmien ominaisuuksia kuin eroja itse toimintojen kehityksessä. Lukuisat menetelmiin liittyvät tekijät, kuten tehtävämateriaalit ja ärsykkeet, tulosmuuttujat, suoritukseen osallistuvat muut kognitiiviset toiminnot ja tehtävien sensitiivisyys, vaikuttavat toiminnanohjauksen kehitystä koskeviin tuloksiin. Toiminnanohjauksen vaikeuksien kliinisessä arvioinnissa uusi kyselymenetelmä soveltui käytettäväksi vaikeuksien seulonnassa sekä koulussa esiin tulevien toiminnanohjauksen profiilien tarkemmassa arvioinnissa.

Acknowledgements

This project has taken many years and involved many people to whom I would like to express my sincerest thanks.

With the deepest gratitude I remember my first supervisor Professor Marit Korkman. She was the supervisor of my internship in the Lastenlinna hospital, my pro graduate and licentiate's theses, and finally my doctoral thesis. Marit truly supported her younger colleagues and generously shared her research ideas and data with others. Her inspiration and wisdom have had a lasting influence on this work as well as on my entire professional being.

I am very grateful to Professor Laura Hokkanen for supervising this work during the last years. Her patient and determined guidance has helped me to finish this work. Laura's constructive advice and positive support made this time more relaxed than it would otherwise have been.

I am most indebted to Pekka Lahti-Nuuttila, M.A., for the statistical guidance and for willingness to devote his time whenever I needed it. In addition to his expertise in statistical methods, Pekka has given me valuable insights into the cognitive aspects of executive functions.

My warm thanks also go to my other co-authors and collaborators. From the Keskittymiskysely team I thank Sari Jämsä, M.A., for her perceptive observations and for the pleasure of working with her during the numerous hours that we spent refining the details of the questionnaire, and Taru Häyrinen, LicPsych, for administering the clinical data collection and for sharing her clinical experience with us. I thankfully remember Pekka Heiskari who inspired us with his enthusiasm for developing new assessment methods and who gave us the opportunity to collaborate with Psykologien Kustannus. My heartfelt thanks go also to Dr. Vesa Närhi for co-authoring the last article and for his profoundly helpful comments on the earlier ones.

I am most grateful to all those who participated in the collection of the research data. From the Lastenlinna hospital I thank Dr. Arja Voutilainen for making the data collection possible and Ulla Järvinen for taking such excellent care of the procedure. I am also grateful to all the school psychologists who took part in the Keskittymiskysely normative data collection, and the psychologists and students who participated in the NEPSY and NEPSY-II standardization studies.

I also wish to thank for the financial support of this work from the Finnish Cultural Foundation, the Ebeneser Foundation, the Arvo and Lea Ylppö Foundation, the Finnish Brain Foundation, the Mannerheim League Research Foundation, the Support Association of the Child Psychiatric Institution of Haukkala, the Finnish Concordia Fund, and University of Helsinki.

I express my gratitude to the pre-examination reviewers Professor H. Gerry Taylor and Professor Astri J. Lundervold for their insightful comments. I am especially delighted and grateful to Professor Timo Ahonen for agreeing to act as the opponent at the public defense of this thesis.

From the University of Helsinki I sincerely thank the Institute of Behavioural Sciences and Professor Jussi Saarinen for admitting me a place to work in during my post-graduate studies. I am grateful to Professor emeritus Juhani Vilkki for kindly writing me recommendations. I also warmly thank Outi, Johanna, and Anu for sharing the ups and downs of doctoral studies, and Hely, Maarit, and Sanna for the possibility to work in such a friendly atmosphere.

During all these years the Neuropsychological Rehabilitation Center Larmis has provided me a nest for both professional work and recreation. I thank each of my colleagues, past and present, for the collaboration, the debates, and the fun that we have had.

Finally, my wholehearted thanks go to my friends and my family. Especially and with all my heart I thank Vesa for sharing the diversity of life with me and for bringing Fiina and Aapo into my life. My loving thanks go to my daughters Anna and Iina, for being who they are, for growing and developing, and for giving me the chance to develop with them.

List of original publications

This thesis is based on the following original publications, referred to in the text by Roman numerals I–III.

I 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*, 407–428. doi:10.1207/S15326942DN2001_6

II Klenberg, L., Närhi, V., Korkman, M., & Hokkanen, L. (2014). Examining methodological variation in response inhibition: The effects of outcome measures and task characteristics on age-related differences. *Child Neuropsychology*. Aug 30:1–17. [Epub ahead of print] doi:10.1080/09297049.2014.950215

III Klenberg, L., Jämsä, S., Häyrinen, T., Lahti-Nuuttila, P., & Korkman, M. (2010). The attention and executive function rating inventory (ATTEX): Psychometric properties and clinical utility in diagnosing ADHD subtypes. *Scandinavian Journal of Psychology, 51*, 439–448. doi:10.1111/j.1467-9450.2010.00812.x

The articles are reprinted with the kind permission of the copyright holders. In addition, some unpublished data are presented.

Abbreviations

η_p^2	partial eta squared (effect size)
ADHD	attention deficit disorder
ADHD-C	attention deficit disorder, combined subtype
ADHD-H	attention deficit disorder, predominantly hyperactive-impulsive subtype
ADHD-I	attention deficit disorder, predominantly inattentive subtype
ADHD RS-IV	ADHD Rating Scale-IV
ANCOVA	analysis of covariance
ANOVA	analysis of variance
ATTEX	Attention and Executive Function Rating Inventory ATTEX
AUC	Area Under the Curve (ROC analysis)
BDEFS-CA	Barkley Deficits in Executive Function Scale—Children and Adolescents
BRIEF	Behavior Rating Inventory of Executive Function
CEFI	Comprehensive Executive Function Inventory
CHEXI	Childhood Executive Function Inventory
DEX-C	Dysexecutive Questionnaire for Children
D-REF	Delis Rating of Executive Functions
EF	executive functions
FTF	Five-to-Fifteen Questionnaire
MANCOVA	multivariate analysis of covariance
MANOVA	multivariate analysis of variance
NEPSY	Developmental Neuropsychological Assessment NEPSY
NEPSY-II	Developmental Neuropsychological Assessment NEPSY-II
ROC	Receiver Operating Characteristic Analyses
SD	standard deviation
WCST	Wisconsin Card Sorting Test

1 Introduction

Executive functions (EFs) are essential for self-regulation and all voluntary actions. For children, EFs play an important role in learning and functioning in school environments. Developmental difficulties in EFs are common in many childhood disorders and especially in attention deficit disorders (ADHD). Knowledge of the normative age-related differences is needed before individual differences can be examined. In spite of extensive research, however, the developmental findings remain unclear due to methodological challenges in measuring EFs. In the clinical assessment of EF difficulties, EF measures that are sensitive to every-day difficulties are required.

1.1 Definitions of executive functions

Executive functions refer to cognitive abilities responsible for directing, controlling, and coordinating other cognitive functions and behavior. The concept originated from research on patients with frontal lobe damage, and the theoretical models of EFs are based on the specific task performance that differentiates frontal lobe patients from control participants. EFs are involved in all purposeful activity, but they are especially needed for in novel, non-routine situations, or during complex problem solving. In the ever-changing and complex everyday environments, EFs appear in behaviors such as waiting for turns, paying attention, and shifting from one action to another, and they enable us to create a plan before taking action, to start acting according to the plan without external support, and to stay on task until its completion.

At present, there is no unifying theory of EFs, and specifications of the component processes vary across models. Theories based on research on frontal lobe dysfunction typically include the processes of anticipation, goal selection, planning, organization, initiation, monitoring, and inhibition as components of EFs (Lezak, 1993; Luria, 1973; Stuss & Benson, 1986). Working memory, the ability to actively hold and process information in mind, has also been conceptualized as a component included in or affecting EFs (Baddeley, 1996). The probably most widely adopted contemporary model of EFs has been introduced by Miyake et al. (2000). Using selected tasks, they distinguished three basic EFs of inhibition, working memory, and shifting between tasks or mental sets. According to the hierarchical model, these basic functions contribute differentially to performance on more complex tasks of planning, flexibility, and strategy use. In developmental samples, the model of Miyake et al. (2000) has been examined with mixed results. Several recent studies indicate that the EFs are less separable in preschool and school-age children, and the correlations among EF tasks are larger in children than in adults (Brydges, Fox, Reid, & Anderson, 2014; Lee, Bull, & Ho, 2013; Wiebe et al., 2011). Miyake et al. have suggested that the shared variance among EF tasks may arise from inhibitory control and/or working memory. Similarly, Barkley (1997, 2006), in his model of EF dysfunction in children with ADHD, has postulated inhibition, and later also working memory, as the foundation for other EFs.

Inhibition, in itself, is a multidimensional concept. There exists several different models of the processes included in inhibition, and the names and meanings of these processes vary across models (Barkley, 1997; Friedman & Miyake, 2004, Kipp, 2005; J. T. Nigg, 2000). Response inhibition, also defined as motor inhibition or behavioral inhibition, has been included in all models. It refers to the ability to intentionally suppress automatic, prepotent, or overlearned responses. Other often postulated inhibitory processes include interference control, the ability to resist distraction from the external environment, and cognitive control, the ability to suppress disruption by irrelevant thoughts and ideas (Friedman & Miyake, 2004).

Although the processes of attention are not commonly included in the models of EFs, they are in several ways overlapping with them. Luria (1973; 1980) used the concept of voluntary attention to describe the directivity and selectivity of perception and actions, thus emphasizing the regulative role of attention in goal-directed behavior. Current, modular models of attention often include executive attention as a separable process responsible for coordination of the focusing, sustaining, and shifting of attention (Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991; Posner & Rothbart, 1998; Shallice, 1982; van Zomeren & Brouwer, 1994). On the other hand, the core attention system can be seen as the foundation for EFs. Developmental models suggest that the ability to focus attention and voluntarily select stimuli are necessary prerequisites for the development of EFs (Garon, Bryson, & Smith, 2008; Posner, Sheese, Odludaş, & Tang, 2006).

In summary, EFs include a large group of partly separable and also overlapping cognitive processes and behaviors. According to current models and developmental studies, the processes of inhibition and working memory can be postulated as the relatively simple core processes of EFs. The relatively more complex EFs include goal-oriented behaviors such as initiation, planning, monitoring, and evaluating actions. Processes of focusing, shifting, and sustaining attention are closely related to EFs and enable the selection of adequate information for the performance of EFs.

1.2 Assessment of executive functions

1.2.1 Performance based measures

Performance based tasks and tests aim to give structured and standardized information of the EFs. Traditional measures include tests involving complex EFs such as planning (e.g., Tower of London; Shallice, 1982), shifting of response set (e.g., Wisconsin Card Sorting Test [WCST]; Grant & Berg, 1948), and fluency (e.g., Controlled Oral Word Association; Benton & Hamsher, 1989), as well as relatively more simple tasks of inhibition (e.g., Stroop task; Stroop, 1935), attention (e.g., Continuous Performance Test, Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956), and working memory (e.g., Digit Span Backward; Wechsler, 1955). Derived from studies on adult brain damage or normal cognitive performance, the original tasks are intended for assessment of EFs in adult populations.

EF measures appropriate for children are either modifications of the adult tasks or specifically designed for use with children. Because developmentally suitable measures are required for a wide age range, the variety of child-appropriate EF measures is extensive. The measures of response inhibition, for example, range from simple motor inhibition tasks, such as the Statue (Korkman, Kirk, & Kemp, 1997, 1998), to stop signal tasks, go-no go tasks, and modifications of the Stroop Task. Moreover, to obtain sensitivity to different age groups, several different versions of each task have been developed. There exists various versions of most EF tasks, e.g, tower (Borys, Spitz, & Dorans, 1982; Korkman et al., 1997, 1998; Levin et al., 1996), fluency (Korkman et al., 1997, 1998; Regard, Srauss, & Knapp, 1982; Welsh, Pennington, & Groisser, 1991), continuous performance (Conners, 1995; Keith, 1994; Korkman et al., 1997, 1998), and Stroop-like tasks (Carlson & Moses, 2001; Delis, Kaplan, & Kramer, 2001; Gerstadt, Hong, & Diamond, 1994; Korkman et al., 1997, 1998).

The Stroop task is perhaps the most widely used and researched measure of inhibition (e.g., Ben-David, Nguyen, & van Lieshout, 2011; Dimoska-Di Marco, McDonald, Kelly, Tate, & Johnstone, 2011; Lansbergen, Kenemans, & van Engeland, 2007; Schwartz & Verhaeghen, 2008), and there are several different child-appropriate modifications of the task. The original Stroop Color-Word Task (Stroop, 1935) was intended for measuring interference control and consisted of tasks measuring reading speed (reading color words printed in black) and naming speed (naming the color of squares), and two inhibition tasks (reading color words printed in incongruent color or naming the incongruent color of ink of color words). An interference score was obtained for the pair of color naming tasks, shown as the difference between the incongruent naming task and the basic naming task.

The reading requirement of the Stroop task sets obvious limits to young children and also for some school-age children. If reading is not automatized, the task is not a valid measure of inhibition (van Mourik, Oosterlaan, & Sergeant, 2005). Modifications for children are typically pictorial tasks of naming pictures (e.g., the Day-Night task; (Gerstadt et al., 1994). In an incongruent naming task, the child is to inhibit a prepotent naming response (e.g., naming a picture of sun as “Day”) and to give a new, rule-based naming response instead (e.g., naming a picture of sun as “Night”). For older children, more complex modifications may also involve tasks that require switching between several different rules for naming (e.g., Huizinga, Dolan, & van der Molen, 2006). In these child-appropriate modifications, the task materials in themselves do not include conflicting information, and, thus, these Stroop-like tasks are more “pure” measures of response inhibition than measures of interference control.

1.2.2 Behavioral rating scales

Behavioral measures assess EFs in everyday environments. Rating scales provide structured information on the child's EF behaviors in the daily life. This information can be obtained from multiple environments, e.g., home and school, and from multiple respondents, e.g., parents and teachers. Behavioral EF measures are especially needed for in clinical assessment, as they help to identify and allocate support for children who suffer from EF difficulties in everyday situations.

Recently, several rating scales designed for the evaluation of EF behaviors in children have been published in the United States and Europe. The most well-known and widely used EF rating scale is the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). Other, more recently published rating scales include the Dysexecutive Questionnaire for Children (DEX-C; Emslie, Wilson, Burden, Ninno-Smith, & Wilson, 2003), Childhood Executive Function Inventory (CHEXI; Thorell & Nyberg, 2008), the Barkley Deficits in Executive Function Scale - Children and Adolescents (BDEFS-CA, Barkley, 2012), the Delis Rating of Executive Functions (D-REF; Delis, 2012), and the Comprehensive Executive Function Inventory (CEFI, Naglieri & Goldstein, 2013). EF rating scales typically include several scales representing different components of EFs (Table 1). The selection of items and scales, however, varies across rating scales according to the clinical aims and theoretical foundations. For clinical use in Finland, adequate local normative or psychometric data for the BRIEF or any other EF rating scale have not been available. The Five-to-Fifteen questionnaire (FTF; Kadesjö et al., 2004; Korkman, Jaakkola, Ahlroth, Pesonen, & Turunen, 2004) with Nordic normative data has been in use for parent evaluation of EF difficulties, but a comprehensive rating scale for assessment of EFs in school situations has been lacking.

Table 1. EF rating scales

Rating scale	Included EF scales
Behavior Rating Inventory of Executive Function (BRIEF) (Gioia et al., 2000)	Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, Monitor
Dysexecutive Questionnaire for Children (DEX-C) (Emslie et al., 2003)	Emotional/Personality, Motivational, Behavioral, Cognitive
Five-to-Fifteen (FTF) (Kadesjö et al., 2004)	Attention, Hyperactive-Impulsive, Hypoactive, Planning and Organizing
Childhood Executive Function Inventory (CHEXI) (Thorell & Nyberg, 2008)	Working Memory, Planning, Inhibition, Regulation
Barkley Deficits in Executive Function Scale - Children and Adolescents (BDEFS-CA) (Barkley, 2012)	Self-Management to Time, Self-Organization/ Problem Solving, Self-Restraint, Self-Motivation, Self-Regulation
Delis Rating of Executive Functions (D-REF) (Delis, 2012)	Attention/Working Memory, Activity Level/Impulse Control, Compliance/ Anger Management, Abstract Thinking/Problem-Solving
Comprehensive Executive Function Inventory (CEFI) (Naglieri & Goldstein, 2013)	Attention, Emotion Regulation, Flexibility, Inhibitory Control, Initiation, Organization, Planning, Self-Monitoring, Working Memory

1.2.3 Methodological issues in the assessment of executive functions

Several methodological issues complicate the assessment of EFs. For the performance based tasks and tests, the major methodological challenges relate to the multidimensionality and the large variety of EF tasks (e.g., Friedman & Miyake, 2004; Jurado & Rosselli, 2007; Miyake et al., 2000). All EF tasks are multidimensional in that they involve multiple cognitive processes other than the targeted EFs. For example, the Stroop tasks measure interference control of distracting stimuli or inhibition of prepotent responses, but they also involve perceptual, language, and motor processes as well as processing speed and working memory. Individual differences in task performance may therefor reflect change in any of these processes.

For some EF tasks, there exist specific methods for cancelling out the effects of other processes. With the Stroop task, the original method was to calculate a difference score by subtracting the performance time of the basic task from the performance time of the incongruent task (Stroop, 1935). Several studies have shown, however, that the difference score may be sensitive to group differences in the basic task processing speed, slower individuals having larger inhibition scores than faster individuals (Ben-David et al., 2011; Lansbergen et al., 2007; Schwartz & Verhaeghen, 2008). Alternative methods for controlling the effects of non-inhibitory processes include calculation of a ratio score by dividing the time for the incongruent task with the time for the control task, and the analysis of covariance (ANCOVA), with the score from the basic naming task used as a covariate in the analysis of group effects on the incongruent task. The different controlling methods have been compared mainly in studies on adults. Studies with preschool age children have typically left the effects of naming uncontrolled, as the performance in basic naming is not recorded. Studies that include school-age children have usually applied either the difference score (Leon-Carrion, García-Orza, & Perez-Santamaría, 2004; Prencipe et al., 2011) or ANCOVA (Huizinga et al., 2006), and only recently also the ratio score (Ikeda, Okuzumi, & Kokubun, 2013; Macdonald, Beauchamp, Crigan, & Anderson, 2014).

The large variety of EF tasks, although necessary for appropriate measurement of performance on different developmental levels, sets another methodological challenge. Different tasks may intend to measure similar EF processes, but differences in task requirements, stimuli, and outcome variables make comparisons across tasks difficult. For example, the tasks intended to measures response inhibition all involve withholding a dominant, prepotent, or automatic response. There are, however, considerable differences in the level of difficulty of the tasks (prepotency of inhibited response, level of conflict between the prepotent and rule-based response) and/or in the involvement of other cognitive processes (e.g., working memory and motor processes). The use of different outcome variables (time *vs.* errors) further complicates across-studies comparisons.

A clinically significant limitation of EF tasks is that performance on these tasks may not reflect problems that are evident in real life (Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Gioia, Isquith, Kenworthy, & Barton, 2002; Jurado & Rosselli, 2007). Tests are usually highly structured, and the examiner is the one to determine how and when the task is to be completed, thus leaving very little need for EFs in completing the task. As a result, a person who performs adequately on EF tasks may still have substantial difficulties in real-life EFs. In line with clinical observations, the correlations between EF tests and everyday EFs have repeatedly been shown to be low or moderate

(Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Bennett, Ong, & Ponsford, 2005; Lawrence et al., 2004).

The methodological challenges of behavioral EF measures relate to the inherent subjectivity of perceptions of the respondent and the psychometric properties of measures. The rater's perceptions of EF behaviors may reflect trait effects (the actual child behavior that is consistent across settings), source effects (child behavior that differs across settings and/or factors related to the rater), and measurement error (DuPaul, 2003). Studies comparing parent and teacher ratings often show low correlations between the respondents (Jarratt, Riccio, & Siekierski, 2005; Mares, McLuckie, Schwartz, & Saini, 2007; Sullivan & Riccio, 2007). This indicates large source effects that may arise from differences in the child behavior across settings or from factors related to the rater, e.g., the history of interaction with the child or former experience of children with problem behaviors. In examining ratings of ADHD symptoms, Gomez, Burns, Walsh, & Moura (2003) found little effect of measurement error, but greater source effect than trait effect on ADHD ratings. In a later study, however, Gomez (2007) showed that the source effects were associated more strongly with situation specificity of behavior than with biased perceptions of the raters. According to these studies, behavioral ratings are sensitive to effects of situation on the child's behavior and, thus, they may be especially useful when gathering information of variation in behavior or of behaviors typical for a specific environment.

The problem of measurement error is related to the psychometric properties of rating scales. The inadequate data of psychometric properties are often critical to the scales functioning (Collett, Ohan, & Myers, 2003; Myers & Winters, 2002). The scale structure and ability to detect behavioral changes over time should be evidenced with reliability estimates. The scales theoretical coherence and ability to discriminate children with and without EF difficulties should be confirmed with validity studies. For clinical application of rating scales, regionally applicable and up-to-date normative data are also necessary.

To summarize, the performance based and behavioral measures apparently tap different aspects of EFs. The performance based tasks give information of specific cognitive EFs that appear in a structured context. This performance may be affected by other cognitive or motor processes. There exists a multitude of EF task modifications for children, and differences in task characteristics and inconsistent use of outcome variables hamper comparisons across studies. Questionnaires and rating scales provide information of EF behaviors in real-world situations from the perspective of multiple observers and multiple environments. The lack of adequate psychometric properties and normative data often reduce usability of rating scales.

1.3 Development of executive functions

The first elements of EFs emerge in naturalistic settings. When playing with toys, 6- to 8-month-old infants are able to remember where a toy is hidden over a short delay, to focus attention towards an interesting toy while resisting distracting effects of other toys, and to stop playing in response to adult request (Diamond, 1985; Garon et al., 2008; Goldman-Rakic, 1987; Kochanska, Tjebkes, & Forman, 1998). In later childhood and adolescence, the developmental change has been explored using EF tasks. Age-related improvement in EF task performance has been shown to continue until adulthood (Huizinga et al., 2006; Welsh et al., 1991).

A few studies have examined the development of EFs by including a wide age-range of children and a large array of different EF tasks. In the classic study of Welsh et al. (1991), a cross-sectional sample of 3- to 12-year-old children and an adult group were assessed with a battery of EF tasks selected from neuropsychological and developmental psychology literature. Children reached the level of the adult group performance first in two simple tasks of strategic planning, a visual search task and a modified tower test, at the age of 6 years. Performance on a more complex visual search task and on the WCST reached the adult level by the age of 10 years. The latest to show developmental change were a verbal fluency task and a motor sequencing task, in which performance reached the adult level by the age of 12 years. In a more recent study, Huizinga et al. (2006) included four age groups of 7-, 11-, 15-, and 21-year old participants and used computerized EF tasks of inhibition, working memory, and shifting. Latent components of the measures and a control measure of processing speed were used as outcome variables for EF development. The obtained developmental pattern indicated that inhibition and shifting reached mature levels by adolescence, while working memory and basic processing speed followed a more protracted course of development into young adulthood.

A more typical approach has been to focus on development of selected EF processes in a narrow age range of children. Studies on development of attention have shown that the ability for joint attention emerges very early, already in the first months of life, and the development of voluntary focused attention continues between ages from 1 to 4 years (Rueda, Posner, & Rothbart, 2005; Ruff & Capozzoli, 2003). Sustaining the focused attention develops for a long time starting from infancy and continuing at least until the age of 10 years (Garon et al., 2008; Rebok et al., 1997).

Response inhibition is perhaps the most investigated EF process. Elements of response inhibition, e.g., the ability to voluntarily stop an activity, have been shown to emerge in the first year of life (Garon et al., 2008; Kochanska et al., 1998). During the preschool period, developmental change has been examined using simple inhibition tasks. For example, in conflict tasks that require inhibition of an automatic response and performance of another, conflicting response (e.g., the Hand Game; Hughes, 1998), significant improvement has been reported between the ages from 3 to 5 years (Garon et al., 2008; Watson & Bell, 2013). In pictorial Stroop tasks, age-related differences have been reported from 3 years (Carlson, 2005) and continuing at least until 8 years of age in computerized versions of the task (McAuley, Christ, & White, 2011). Further development in school age has been reported in studies using the classic Stroop task (Hummer, 2011; Leon-Carrion et al., 2004; Prencipe et al., 2011; Wu et al., 2011) or more complex versions of non-reading Stroop tasks. Huizinga et al. (2006) used a computerized task that required the participants to first name the color or the orientation of a smiley figure, and then, in an interference trial, to respond according to two different new rules (responding to the color according the orientation of the figure). In this complex task the developmental change continued until 21 years.

Very similarly to attention and response inhibition, the other EFs have shown early emerging but protracted developmental improvement. For example, elements of working memory appear already in 6 month old infants, but the length of time for holding information in mind, and the number of items held in mind, increase over and beyond the preschool years (Garon et al., 2008). The more complex processes of updating and manipulating working memory representations emerge in 2-year-old children and improve at least through adolescence (Best & Miller, 2010; Conklin, Luciana, Hooper, & Yarger, 2007; Luciana & Nelson, 1998; McAuley & White, 2011). The ability to shift response set seems to emerge somewhat later than other EFs. Children from 3 to 4 years have some ability to flexibly shift their responses on the basis of environmental feedback, but 5-year-olds still continue to make lots of perseverative errors (Garon et al., 2008). Similarly to the other EFs, the ability to shift between responses continues to develop until adolescence (Davidson, Amso, Anderson, & Diamond, 2006; Huizinga et al., 2006).

In summary, EFs develop for a protracted period, starting from infancy and continuing into adulthood. The different processes of EFs develop gradually, seemingly in a sequential (Jurado & Rosselli, 2007), multistage (Passler, Isaac, & Hynd, 1985; Romine & Reynolds, 2005), or hierarchical process (Garon et al., 2008; Welsh et al., 1991). There is, however, no clear evidence of staging of the development. In all EF processes, significant developmental change has been reported throughout childhood. The methodological challenges related to the variability in the tasks, outcome measures, and included samples, and the multidimensionality of tasks set considerable challenges for studies on EF development. A potential approach for systematic assessment of EFs is to employ tasks that are appropriate for a large age range and use these to study age-related differences in a single sample of children of different ages. Another approach is to employ only one measure and to focus on disentangling the different factors that may affect age-related differences in the task performance.

1.4 Difficulties in executive functions

Difficulties in EFs may manifest themselves in a variety of ways. Children may act impulsively and have difficulty in stopping activity, they may have difficulty in getting started and keeping active, or they may get easily distracted and have difficulty in staying focused. Often, impairments in EFs show up globally, affecting all areas of behavior (Lezak, 1995). However, because EFs are especially needed for in novel situations or when complex problem solving is required, these difficulties may be especially impairing for academic functioning (Best, Miller, & Naglieri, 2011; St Clair-Thompson, & Gathercole, 2006). Learning situations at school set specific demands for the ability to sustain attention, to plan and organize materials, and to monitor and evaluate one's own behavior as well as for behavioral control in the complex social context.

Difficulties in EFs are associated with many childhood disabilities such as learning disorders (Booth, Boyle, & Kelly, 2010; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011), conduct disorder (Herba, Tranah, Rubia, & Yule, 2006; Sergeant, Geurts, & Oosterlaan, 2002), autism spectrum disorders (Hill, 2004), and, especially, with ADHD (Barkley, 1997; Barkley, 2006; Doyle et al., 2005; Nigg, 2001). Of the three ADHD subtypes recognized in the Diagnostic and Statistical Manuals of Mental Disorders (DSM-IV; American Psychiatric Association, 1994; DSM-5; American Psychiatric Association, 2013), children with combined symptoms of inattention and hyperactivity-impulsivity (ADHD-C) and children with predominantly inattentive symptoms (ADHD-I) have shown significant impairment in EFs (e.g., Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Children with predominantly hyperactivity-impulsivity symptoms (ADHD-H) have been included in only a few studies, and this subtype has not been associated with EF impairment (Willcutt et al., 2005).

1.4.1 Executive function difficulties in ADHD

In studies using performance based EF measures, groups of children with ADHD have shown more impairment than controls in most EF tasks and tests (Willcutt et al., 2005). There is, however, considerable variability within the ADHD group, and not all children with ADHD have weaknesses in EF performance (Doyle et al., 2005; Nigg & Casey, 2005; Sonuga-Barke, 2005; Willcutt et al., 2005). To some extent, this may reflect true variability within the ADHD group. On the other hand, inconsistent findings may result from the lack of sensitivity of performance based EF measures (Anderson et al., 2002; Bennett et al., 2005; Biederman, 2008).

According to behavioral ratings, children with ADHD have consistently shown considerable problems in almost all EF behaviors (Gioia et al., 2002; Jarratt et al., 2005; McCandless & O' Laughlin, 2007; Semrud-Clikeman, Walkowiak, Wilkinson, & Butcher, 2010; Sullivan & Riccio, 2007). This is an expected finding, as the EF ratings partly measure similar behaviors as are included in ratings of ADHD symptoms. For example, the EF items related to difficulties in inhibition overlap with ADHD symptoms of impulsivity and motor hyperactivity. The EF rating scales, however, cover a much wider range of different EF behaviors than the symptoms of inattention, impulsivity, and hyperactivity included in ADHD rating scales. Thus, while ADHD rating scales are suitable for diagnostic purposes, rating scales that describe EF behaviors in more detail can potentially give information that is useful when assessing the child's needs for support and planning for interventions.

Those children with ADHD, who have difficulties in EFs, may be at risk for worse developmental outcome (Halperin, Trampush, Miller, & Newcorn, 2008) and especially for worse outcome in academic functioning (Biederman et al., 2004; Biederman et al, 2006). Assessment of EF difficulties that arise in school environments seems thus particularly important. Previous studies examining teacher ratings of behavioral EF difficulties in children with ADHD indicate that the EF problems of children with ADHD are apparent is school environments, but the findings do not consistently show how wide ranging these difficulties are and if there are specific EF profiles typical for children with ADHD (Jarratt et al., 2005; Sullivan & Riccio, 2007; Toplak, Bucciarelli, Jain, & Tannock, 2009). The lack of consistent findings may relate to the variety within the ADHD group, as the previous studies typically have not differentiated between the ADHD subtypes. Only the study of McCandless and O'Laughlin (2007) compared the two subtypes. Using teacher ratings of two index scores and the two scales of inhibition and working memory of the BRIEF, this study found no differences between children in the ADHD-C and ADHD-I groups. Parent ratings, however, have indicated that children with ADHD-C may show larger difficulties in EF behaviors than children with ADHD-I (Gioia et al., 2002; McCandless & O' Laughlin, 2007; Semrud-Clikeman et al., 2010).

In summary, those children with ADHD who have difficulties in EFs have an increased risk for impairment in school settings. Previous studies on teacher rated EF difficulties have indicated elevated scores but no distinct EF profiles for children with ADHD. Rating scales for assessment of school related EF difficulties are necessary for screening and allocating support for these children. In the Finnish clinical practice, a well-studied EF rating scale has been lacking.

2 Aims of the study

This thesis addresses assessment and age-related differences in EFs by using performance based and behavioral measures and by employing three different normative samples. The specific aims of the included three studies are to

1. examine age-related differences in EF components using a large set of performance-based tests in a sample of 3- to 12-year-old children (Study I),

2. examine the effects of outcome measures and task characteristics on age-related differences in response inhibition performance in a sample of 7- to 15-year-old children and adolescents (Study II), and to

3. construct and examine the psychometric properties of a new rating scale for assessment of EF difficulties in children with ADHD (Study III).

3 Methods

The characteristics of participants are summarized in Table 2. In the following sections, the methods of each of the three studies are described in more detail.

Table 2. Summary of the characteristics of participants in Studies I–III

	Number	Age		Male	Mother's education level		
		M	(SD)		Lower	Medium	Higher
Study I							
3-year-olds	40	3.5	(0.3)	52.5 %	27.5 % [a]	40.0 % [a]	30.0 % [a]
4-year-olds	38	4.4	(0.3)	52.6 %	21.5 % [a]	57.9 % [a]	15.8 % [a]
5-year-olds	40	5.4	(0.3)	47.5 %	37.5 % [a]	45.0 % [a]	17.5 % [a]
6-year-olds	40	6.4	(0.3)	47.5 %	27.5 % [a]	45.0 % [a]	27.5 % [a]
7-year-olds	40	7.5	(0.3)	50.0 %	30.0 % [a]	42.5 % [a]	27.5 % [a]
8-year-olds	39	8.5	(0.3)	51.3 %	30.8 % [a]	43.6 % [a]	23.1 % [a]
9-year-olds	40	9.5	(0.3)	47.5 %	42.5 % [a]	32.5 % [a]	17.5 % [a]
10-year-olds	41	10.5	(0.3)	52.5 %	31.7 % [a]	43.9 % [a]	24.4 % [a]
11-year-olds	41	11.5	(0.3)	46.3 %	41.5 % [a]	39.0 % [a]	17.1 % [a]
12-year-olds	41	12.5	(0.3)	52.5 %	39.0 % [a]	29.3 % [a]	31.7 % [a]
Study II							
7-year-olds	80	7.0	(0.1)	56.3 %	3.8 % [b]	53.8 % [b]	42.5 % [b]
9-year-olds	66	9.0	(0.1)	51.5 %	0.0 % [b]	56.9 % [b]	43.1 % [b]
11-year-olds	68	11.0	(0.1)	42.6 %	1.5 % [b]	63.2 % [b]	35.3 % [b]
13-year-olds	62	13.1	(0.1)	41.9 %	1.6 % [b]	67.7 % [b]	30.6 % [b]
15-year-olds	64	15.1	(0.1)	37.5 %	4.7 % [b]	56.3 % [b]	39.1 % [b]
Study III							
Normative	701	10.7	(2.5)	47.8 %	23.1 % [a]	45.1 % [a]	31.8 % [a]
ADHD-C	190	10.3	(2.4)	86.5 %	40.6 % [a]	43.9 % [a]	15.5 % [a]
ADHD-I	25	9.9	(2.5)	84.0 %	8.7 % [a]	65.2 % [a]	26.1 % [a]

[a] Mother's education level: Lower = primary and/or secondary school and/or vocational school (4-12 years); Medium = senior high school or college (12-15 years); Higher = university education (16 years or more).

[b] Mother's education level: Lower = primary and/or secondary school only (4-9 years); Medium = senior high school and/ or vocational school or college (12-14 years); Higher = university education (16 years or more).

3.1 The NEPSY study (Study I)

3.1.1 Participants and procedures

The participants in Study I were children from the Finnish NEPSY (Korkman et al., 1997) standardization study (Table 2). The sample was drawn from day care units and schools in five different localities in Finland during years 1993-1994. 400 typically developing children from ten age groups of 3- to 12-years participated in the study. Children with parent reported diagnosed neurological disorders (e.g., epilepsy, cerebral palsy) or developmental disorders (language, learning, or attention disorder) were excluded from the study.

Approval for the study was granted from municipal authorities in the five different localities. Written informed consent and background information from caregivers were obtained prior to the assessment.

3.1.2 Instrument

Ten subtests from the Finnish version of the Developmental Neuropsychological Assessment NEPSY (Korkman et al., 1997) were used as measures of EFs. The age-appropriate subtests for 3- and 4-year-old children included three subtests, for 5- and 6-year-old children nine subtests, and 7- to 12-year-old children ten subtests.

The *Statue* subtest is a task of response inhibition. The child is to maintain a body position and remain silent, eyes closed, during a 75-sec period, while the examiner tries to provoke responses by producing sounds. Movements and vocalizations are recorded. The score is the number of 5-sec intervals the child is able to stay still, eyes closed, and silent (max score 15).

The *Knock and Tap* subtest is a more complex response inhibition task. First, the child is instructed to knock on the table when the examiner taps on the table, and to tap when the examiner knocks. In the second part, the child is to tap with the side of the fist when the examiner knocks with the knuckles, and vice versa, and not to respond at all when the examiner taps with the palm. The score is the total number of correct responses (max score 30).

The *Auditory Attention* subtest is a continuous performance task where the child responds to the word “red” by putting a red square in a box when listening to a list of 180 words (e.g., “green…now…put…”) on an audiotape. If the response is given within 1 sec, the score is 2, and if given within 2 or 3 sec, the score is 1.

The *Auditory Response Set* subtest is a more complex continuous performance task. This time, the child is to respond to the word “red” by putting a yellow square in the box and to the word “yellow” by putting a red square in the box. When hearing the word “blue”, the child is to put a blue square in the box. Each correct response gives 2 or 1 points, and the total score is the score from correct responses minus the number of incorrect responses (max score 72).

In the *Visual Search* subtest, the child is to locate target pictures (cats) among other figures (e.g., bunnies, houses, trees) as fast as possible. The score is an efficiency index (correct responses minus incorrect responses divided by performance time).

The *Visual Attention* subtest is a cancellation task where the child locates the targets (two pictures of faces) from a linear array of pictures of different faces. The score is an efficiency index (correct responses minus incorrect responses divided by performance time).

The *Tower* subtest is an adaptation of the Tower of London Test (Shallice, 1982) where the child is to move three colored balls to target positions on three pegs in a prescribed number of moves. Only one ball can be moved at a time, a ball cannot be placed elsewhere than on one of the pegs, and only a certain number of balls can be put on each of the different pegs. The score is the number of correctly achieved target positions obtain within the time limit (30-60 sec; max score 20).

In the *Semantic Fluency* subtest, the child first produces as many animal names as possible in 1 min and then produces things to eat and drink in 1 min. The score is the total number of correct words given in the two categories.

In the *Phonemic Fluency* subtest, the child first produces words beginning with the letter “s” as many as possible in 1 min and then produces words beginning with the letter “k” in 1 min. The score is the total number of correct words given in the two categories.

The *Design Fluency* subtest is based on the Five Point Test of Regard (Regard et al., 1982) and the Ruff Figural Fluency Test (Vik & Ruff, 1988). The child draws as many different designs as possible by connecting two or more dots with straight lines in 1 min. In the first part, the dots are in a structured array, and in the second part, the array is more unstructured. The score is the number of correct, different designs produced.

The variables used in Study I differed from those of the standard procedure of the Finnish NEPSY in tasks of auditory attention, visual attention, and verbal fluency, where subscores of tasks instead of combined total scores were used. Additionally, in the visual attention tasks, the variable scores differed from those of the standard procedure. The subtests for different age groups and the variables used in Study I are presented in Table 3.

Table 3. EF tests from the NEPSY used in Study I

	Neuropsychological tests		
3-4 -year-olds	5-6 -year-olds	7-12 -year-olds	Variables used
Statue	Statue	Statue	number of inhibited 5-sec intervals
Knock and Tap	Knock and Tap	Knock and Tap	number of correct responses
	Auditory Attention	Auditory Attention	correct responses - commission errors
	Auditory Response Set	Auditory Response Set	correct responses -commission errors
Visual Search	Visual Search	Visual Search	correct responses - commission errors/ time
	Visual Attention	Visual Attention	correct responses - commission errors/ time
	Tower	Tower	number of correct solutions
Semantic Fluency	Semantic Fluency	Semantic Fluency	number of correct words in the two categories
		Phonemic Fluency	number of correct words in the two categories
	Design Fluency	Design Fluency	number of correct designs

3.2 The NEPSY-II study (Study II)

3.2.1 Participants and procedures

The participants in Study II were from the Finnish NEPSY-II standardization study (Korkman, Kirk, & Kemp, 2008). The NEPSY-II sample was recruited by randomly selecting 6,006 children, aged 3 to 15 years and living in eight selected localities, from the Finnish population register during years 2006-2007. The original standardization sample consisted of 923 volunteered children (15.4% of the invited). In this sample, the distribution of mother’s education level was compared to the distribution of education level for 25- to 54- year-old women in the general population in the year 2007 (Official Statistics of Finland, 2012). The sample differed from the general population in mother's education, $X^2(2) = 59.26$, $p < .001$, the lower and higher education levels being less common in this sample (lower 2.4 %, higher 38.2 %) than in the general population (lower 13.3 %, higher 45.2 %), and the medium education level being more common in this sample (59.1 %) than in the general population (45.2 %). Ethical approval for the study was granted from the Committee for Research Ethics at Åbo Akademi University. Written informed consent and background information from caregivers were obtained prior to the assessment.

Children (n = 398) belonging in the age groups of 7 (82-86 months), 9 (106-110 months), 11 (130-134 months), 13 (154-158 months), and 15 (178-182 months) years were included in Study II (Table 2). Prior to statistical analyses, children with first language other than Finnish (n = 3) and children who according to parent reports had difficulties related to attention, social interaction, learning, or other developmental disabilities (n = 55) were excluded. The final sample consisted of 340 typically developing children and adolescents from the five age groups. There were no significant differences in the distributions of sex, X^2 (4) = 6.73, p = .151, and mother's education level, X^2 (8) = 7.41, p = .494, across the age groups.

Of the 340 children, nine (2.6 % of the participants) had missing values in the Switching tasks, and one had missing values in the Inhibition tasks. All these children had high error rates in the Naming and/or Inhibition tasks. As the percentage of missing values was very small, the Expectation-Maximization estimation, based on existing values in other tasks in the Inhibition subtest, was used to impute values for cases with missing values. The imputed data were used in the analyses.

3.2.2 Instrument

Three tasks from the NEPSY-II Inhibition subtest (Korkman, Kirk, & Kemp, 2007a; Korkman et al., 2008) were used as measures of speeded naming (the basic task), inhibition, and switching inhibition. The tasks included two conditions with different stimuli, the first a series of black and white shapes (circles and squares), and the second a series of black and white arrows (upward and downward). In both conditions, first the task of naming, then inhibition, and finally switching were presented.

The tasks and outcome variables are presented in Table 4. For the inhibition scores, the time and error scores were first combined into corrected time scores using the original Stroop protocol (Stroop, 1935): For each uncorrected error, the participant's average stimulus naming time in the task multiplied by 2 was added to the performance time. As the correlations of corrected time scores between the two conditions were high, ranging from .72 to .82, the corrected time scores of the two conditions were summed into the corrected time score of naming, a corrected time score of inhibition, and a corrected time score of switching. The inhibition and switching scores were calculated using these summed scores. For the inhibition difference score, the corrected time score of naming was subtracted from the corrected time score of inhibition, and for the switching inhibition difference score, the corrected time score of naming was subtracted from the corrected time score of switching. To calculate the ratio scores, the corrected time score of inhibition was divided by the corrected time score of naming, and the corrected time score of switching was divided by the corrected time score of naming. For all scores larger values indicated poorer performance.

Table 4. Tasks from the NEPSY-II Inhibition subtest used in Study II

	Variables used	
Task	Analyses of task performance	Analyses of inhibition scores
Naming		corrected time score
Shapes	total time, number of errors	
Arrows	total time, number of errors	
Inhibition		corrected time score; difference score; ratio score
Shapes	total time, number of errors	
Arrows	total time, number of errors	
Switching		corrected time score; difference score; ratio score
Shapes	total time, number of errors	
Arrows	total time, number of errors	

3.3 The ATTEX study (Study III)

3.3.1 Participants and procedures

The 7- to 15-year-old children and adolescents participated in the Attention and Executive Function Rating Inventory (ATTEX) normative study (Table 2). All participants followed the normal curricula in general education classes. Ethical approval for the study was granted from the Helsinki University Central Hospital Ethical Committee for Pediatrics, Adolescent Medicine, and Psychiatry. In all groups, written informed consent and background information from caregivers were obtained prior to teacher ratings.

The normative group was derived by combining two samples. In both samples, children with a parent reported diagnosis of ADHD-C or ADHD-I were excluded. The first sample was collected during years 2005–2006 from 45 schools in Finland. The attrition rate of this sample is not known, but the teachers' estimation was that there were very few refusals. The second sample was recruited from the Finnish NEPSY-II standardization study, collected during years 2006–2007 (Korkman et al., 2008). Of the total of 923, all participants within the appropriate age range received a request to take part in the ATTEX study. Of 482 rating inventories, 194 (40.3 % of the targeted sample) were returned. No differences between respondents and non-respondents were found according to the child's age, sex, teacher reported learning difficulties, or parent education level.

Children in the ADHD-C and ADHD-I groups were recruited from consecutively examined or followed-up patients from the Outpatient Clinic of Pediatric Neurology of the Helsinki University Hospital, Finland, during years 2005–2007. They were diagnosed by clinically experienced child neurologists or resident doctors according to the DSM-IV criteria for ADHD combined subtype or the predominantly inattentive subtype. The diagnostic evaluation of symptoms of ADHD and co-occurring disorders included a clinical diagnostic interview of the child and parents (developmental history, symptoms related to developmental disorders), parent ratings of the ADHD Rating Scale-IV (ADHD RS-IV, DuPaul, Power, Anastopoulos, & Reid, 1998) and the Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997; Koskelainen, Sourander, & Kaljonen, 2000), and teacher reports (learning history, working habits, and behavior of the child during school days). Children (1) diagnosed with severe physical impairment (e.g., paresis) or neuropsychiatric conditions (e.g., autistic disorder), or (2) following individualized curricula due to general learning disabilities were excluded from the study, but children with co-occurring disorders typical to ADHD were included. The diagnosed co-occurring disorders included specific learning disorders in reading or mathematics (47 in the ADHD-C group, 16 in the ADHD-I group), developmental cognitive disorders in language or motor skills (40 in the ADHD-C group, 4 in the ADHD-I group), disorders of social interaction (15 in the ADHD-C group, 1 in the ADHD-I group), and conduct disorders (11 in the ADHD-C group, 1 in the ADHD-I group). Sixty-nine children (all in the ADHD-C group) were on medication for ADHD. For ethical reasons, medication was not discontinued during the assessment and these children were included in the study.

Group comparisons showed that the proportion of boys was larger in the ADHD-C, $X^2(1) = 90.05$, $p < .001$, and ADHD-I groups, $X^2(1) = 12.67$, $p < .001$, than in the normative group. The distribution of mother's education was also different, the lower education level being more common in the ADHD-C group than in the normative group, $X^2(2) = 31.90$, $p < .001$. Of the co-occurring disorders, learning disorders were significantly more frequent in the ADHD-I group than in the ADHD-C group, $X^2(1) = 15.86$, $p < .001$.

Of the original data (n =923), 48 cases (5.2%) had one or more missing items in the ATTEX. These were replaced with the participant's mean value of the respective scale items. Seven cases (three in the normative, three in the ADHD-C, and one in the ADHD-I group) with more than two missing observations were excluded from data analyses. The analyses were run on both the original and completed datasets. As the results did not differ, the results were based on the imputed dataset.

3.3.2 Instrument

The Attention and Executive Function Rating Inventory ATTEX consists of items that are constructed to describe EF behaviours relevant in school situations. The selection of ATTEX items was based on an integration of EF models and theories to cover the processes of inhibition, attentional control, and regulation of action (Barkley, 1997; Lezak, 1995; Luria, 1973; Mirsky et al., 1991; Stuss & Benson, 1986). Originally, 58 items were selected on the basis of theoretical models, literature review, and clinical experience of the authors. In a preliminary pilot study, the items were further improved and rephrased according to teacher feedback. In a second pilot study, a small sample (N = 43) of ADHD-C-diagnosed children aged 6–16 years (M = 10.8; SD = 2.91; 88.1 % boys) was examined. The analysis of the pilot study data showed weak item-test correlations for three items which were eliminated from the final rating inventory, resulting in 55 items. To enhance clinical utility of the ATTEX especially for intervention planning, a list describing strengths of the child and four open questions related to situational variability of behaviour were also included the questionnaire.

The 55 ATTEX items were grouped into ten scales based on their content and the scales were named following the theoretical models (Table 5). The individual items were scored on a scale of 0 (not a problem), 1 (sometimes a problem), or 2 (often a problem). The original Finnish version of the ATTEX (Klenberg, Jämsä, Häyrinen, & Korkman, 2010) has been published and is available for clinical use in Finland. The English version has been published as an Appendix in the original publication (Klenberg, Jämsä, Häyrinen, Lahti-Nuuttila, & Korkman, 2010). The ATTEX summary total score and summary scale scores were used in the analyses of Study III.

The ADHD Rating Scale-IV: School Version (DuPaul et al., 1998) is a DSM IV-based rating scale that comprises the 18 diagnostic symptoms of ADHD. It has solid psychometric properties and is widely used in diagnostic assessment of ADHD. In Study III, the ADHD RS-IV School Version was used as a criterion measure for the validity assessment of ATTEX.

Table 5. The EF ratings used in Study III

Rated scales	Number of items	Maximum score
ATTEX EF scales	55	110
Distractibility	4	8
Impulsivity	9	18
Motor hyperactivity	7	14
Directing attention	5	10
Sustaining attention	6	12
Shifting attention	4	8
Initiative	5	10
Planning	4	8
Execution of action	8	16
Evaluation	3	6
ADHD RS-IV scales	18	54
Inattention	9	27
Hyperactivity-Impulsivity	9	27

3.3 Data analyses

The main statistical methods of the studies were univariate and multivariate analyses of variance (ANOVA, MANOVA) and univariate and multivariate analyses of covariance (ANCOVA, MANCOVA).

In Study I, separate ANOVAs, with age, sex, and parent education level as between-subjects variables, were performed on each subtest. In the original article, significant results for the separate analyses of each subtest were reported. In the present thesis, a Bonferroni corrected significance level ($p > .005$), correcting for the multiple comparisons, was applied instead. The effects of age were further analyzed using Bonferroni corrected comparisons between each age group and consecutive older age groups. Explorative Maximum Likelihood factor analysis with orthogonal rotations was performed to examine the construct validity of measures.

In Study II, ANOVAs with stimulus condition (shapes and arrows) as a within-subjects variable and age as a between-subjects variable were performed on task time variables. The effects of age were further analyzed by performing trend analyses with polynomial contrasts, and pair-wise contrasts between 7- to 13-year-old groups and the 15-year-old group (Bonferroni corrected significance level $p < .0125$). For condition x age interaction effects, separate ANOVAs for each age group were performed (Bonferroni corrected significance level $p < .01$). For the error variables, separate nonparametric tests for the effects of age (independent samples Kruskal-Wallis Tests with Dunn's Tests) and stimulus condition (related-samples Wilcoxon Signed-Rank Tests, Bonferroni corrected significance levels $p < .0125$) were performed. For the difference and ratio scores, ANOVAs with condition (inhibition *vs.* switching inhibition) as a within-subjects variable and age as a between-subjects variable were performed. Age-related differences were further analyzed with separate univariate ANOVAs for each inhibition score. In all analyses, the effects of sex were controlled by including the categorical variable in the model.

In Study III, the effects of age and sex in the normative group were examined with MANOVAs and follow-up univariate ANOVAs. ATTEX internal consistency was examined using Cronbach's alpha and the criterion validity was evaluated using Pearson's product-moment correlations between ATTEX and ADHD RS-IV scales. Receiver Operating Characteristic (ROC) analyses were performed for the normative *vs.* combined ADHD-groups (separately for girls and boys) and for the ADHD-C *vs.* ADHD-I groups. In the present thesis, additional analyses comparing the EF profiles of ADHD-C, ADHD-I, and normative groups were performed with ANCOVAs and MANCOVAs controlling for the effects of age, sex, and parent education. The contrasts for pairwise comparisons of groups were examined applying the Bonferroni corrected significance level $p < .0167$. The effects of disorders co-occurring with ADHD were also examined with separate ANCOVAs and MANCOVAs.

The effect sizes were measured in Studies II and III using the partial eta squared (η_p^2), values < .06 indicating small effects, .06 - .13 indicating medium effects, and ≥ .14 indicating large effects (Cohen, 1988).

4 Results

4.1 Age-related differences in executive function components

All ten EF subtests of the NEPSY were significantly affected by age (Figure 1).The results from separate ANOVAs are reported in the original article (Klenberg et al., 2001, p. 414–418). In the present thesis, the multiple comparisons in the ten separate ANOVAs were taken into account, and the following effects of age, sex, and parent education remained significant.

The effect of age remained significant in all EF subtests. The development was considered to have reached the 12-year-old level when the performance of a specific age group did not differ from that of any of the older groups. In the Statue subtest, $F(9, 317) = 24.87$, $p < .001$, the 12-year-old level was reached at the age of 6 years, in the Knock and Tap subtest, $F(7, 254) = 5.25$, $p < .001$, at 7 years, and in the Tower subtest, $F(7, 267) = 16.46$, $p < .001$, at 8 years. In the subtests Visual Search, $F(9, 323) = 43.78$, $p < .001$, Visual Attention, $F(7, 254) = 32.18$, $p <.001$, Auditory Attention, $F(7, 260) = 17.51$, $p < .001$, Auditory Response Set, $F(7, 254) = 33.80$, $p < .001$, and Semantic Fluency, $F(9, 331) = 73.61$, $p < .001$, this level was reached at 10 years. In the Phonemic Fluency, $F(5, 201) = 20.76$, $p < .001$, and Design Fluency, $F(7, 263) = 38.78$, $p < .001$, the 12-year-old level was reached at the age of 11 years.

The effect of sex remained significant in the Phonemic Fluency subtest, $F(1, 201) = 32.89$, $p < .001$, girls performing better than boys. In the Statue subtest, a significant interaction between age and sex, $F(9, 317) = 2.97$, $p = .002$, showed that girls performed better than boys at ages from 3 to 5 years, but from 6 years onward the difference disappeared.

The effect of parent education level remained significant only in the Semantic Fluency subtest, $F(2, 331) = 6.80$, $p = .001$, children with higher parent education level (university education) performing better than children with lower parent education level (primary and/ or vocational education).

The inter-correlations and possible latent factors behind the subtests were analyzed using the data of 7- to 12-year-old children. A four-factor solution, with eigenvalues

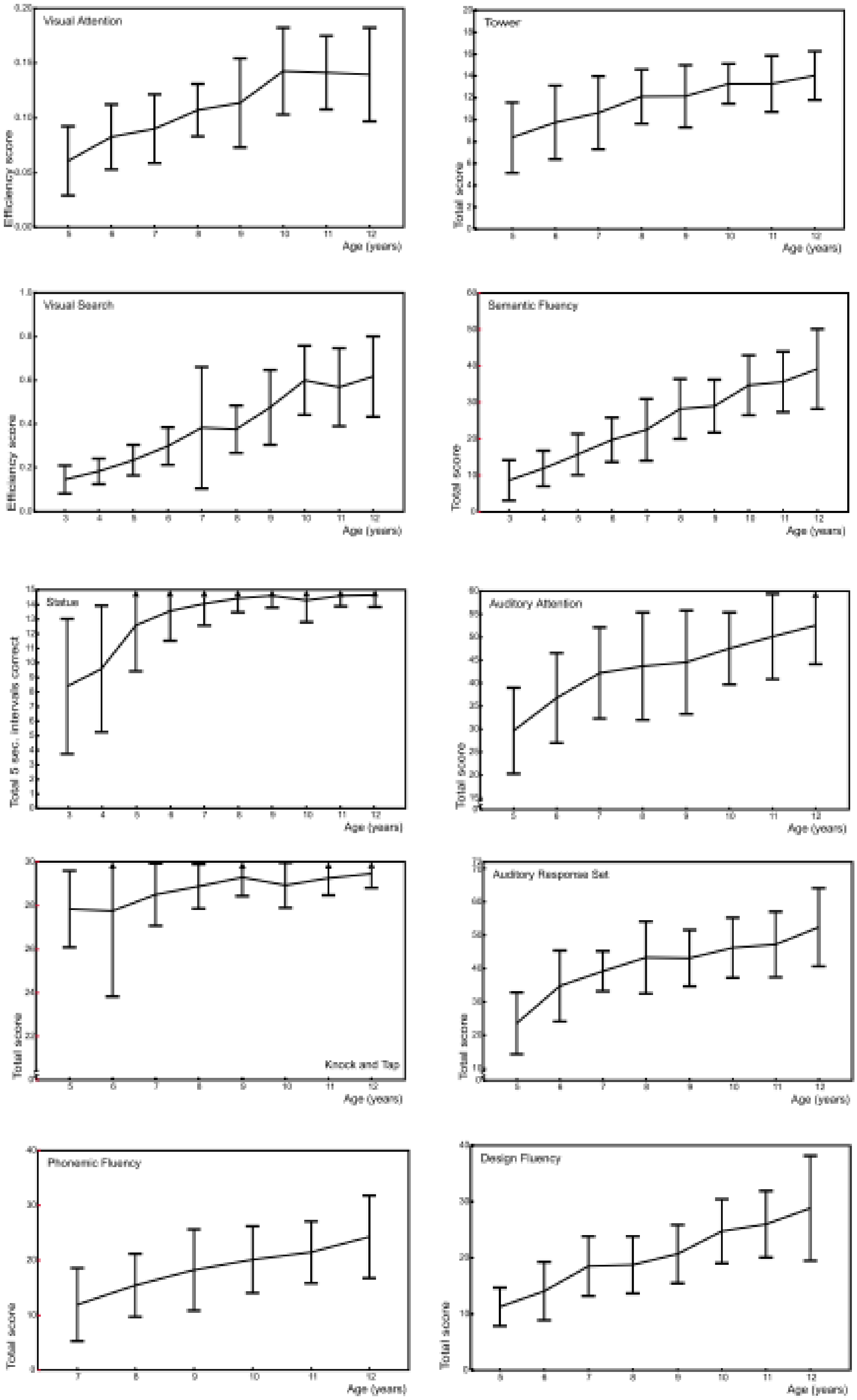

Figure 1. Subtest means and standard deviations (± 1 SD) for age groups.

greater than one, accounted for 56.7% of the total variance (factor loadings are presented in Klenberg et al., 2001, p. 419). Factor 1 included ($r \geq .40$) the subtests Semantic Fluency, Phonemic Fluency, Design Fluency, and Auditory Response Set. The Auditory Response Set subtest, however, had a stronger loading on the third factor. Factor 2 included the subtest of Visual Search, and the Visual Attention subtest also had a weaker correlation (.34) with this factor. Factor 3 included the subtests Auditory Attention and Auditory Response Set. Factor 4 included the Statue subtest. The subtests Tower, Knock and Tap, and Visual Attention did not correlate strongly with the factors in the four-factor solution.

In summary, the results of Study I show that age-related change in EF performance continues at least until the age of 12 year, and the improvement starts to decelerate at different times in different tasks. Using the ten tasks from the NEPSY, the development seemed to proceed from inhibition to attention control, and further to fluency. The exploratory factor structure also implicated relative separateness of performance in these tasks.

4.2 Age-related differences in response inhibition

In the NEPSY-II Inhibition subtest, the analyses of total time scores (Table 6) revealed significant effects of age in all tasks of naming, $F(4, 330) = 153.38$, $p < .001$, $\eta_p^2 = .65$, inhibition, $F(4, 330) = 145.77$, $p < .001$, $\eta_p^2 = .64$; and switching, $F(4, 330) = 132.32$, $p < .001$, $\eta_p^2 = .62$. The linear and quadratic trends of age were significant in all tasks and pairwise comparisons showed that the total times for 7-, 9-, and 11-year-olds were significantly ($p < .001$) longer than the total times for 15-year-olds. The effect of stimulus condition was also significant in all tasks. In the naming tasks, the main effect showed that total times in the Arrows condition were longer than in the Shapes condition, $F(1, 330) = 96.86$, $p < .001$, $\eta_p^2 = .23$, and the interaction effect between stimulus condition and age indicated that the effect was larger for 7- and 9-year-olds than for 11-, 13-, and 15-year-olds, $F(4, 330) = 12.28$, $p < .001$, $\eta_p^2 = .13$. In the inhibition tasks, the main effect, $F(1, 330) = 258.92$, $p < .001$, $\eta_p^2 = .44$, and the interaction effect, $F(4, 330) = 4.07$, $p = .003$, $\eta_p^2 = .05$, similarly showed that the total times in Arrows were significantly longer than in the Shapes for all age groups. In the switching tasks, the main effect of stimulus condition showed a reversed trend, the total times in Shapes being longer than in Arrows, $F(1, 330) = 15.31$, $p < .001$, $\eta_p^2 = .04$.

The analyses of error scores (Table 7) revealed significant effects of age in the Naming Arrows, $H(4) = 26.25$, $p < .001$, Inhibition Shapes, $H(4) = 45.59$, $p < .001$, Inhibition Arrows, $H(4) = 25.26$, $p <. 001$, Switching Shapes, $H(4) = 26.82$, $p < .001$, and Switching Arrows tasks, $H(4) = 47.39$, $p < .001$, but not in the Naming Shapes task. In the Naming Arrows and both inhibition tasks, only the 7-year-olds made more errors than 15-year-olds. In both switching tasks, 7-, 9-, and 11-year-olds made more errors than 15-year-olds. The effect of condition for error scores was significant in the naming tasks, $W = 2\ 387$, $p = .002$, $r = .16$, and inhibition tasks, $W = 9\ 118$, $p < .001$, $r = .29$, the error scores in the Arrows conditions being larger than in the Shapes conditions.

Table 6. Means (*M*) and standard deviations (*SD*) of total time scores for the tasks

	7 years	9 years	11 years	13 years	15 years
Task	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Naming Shapes	33.1 (6.6)	25.9 (4.3)	22.7 (4.0)	19.7 (2.5)	18.9 (3.2)
Naming Arrows	38.7 (9.3)	30.1 (5.7)	24.0 (3.8)	20.8 (3.2)	20.0 (3.7)
Inhibition Shapes	47.1 (12.5)	33.6 (6.7)	28.4 (5.2)	23.9 (3.4)	21.9 (4.0)
Inhibition Arrows	56.8 (15.2)	42.3 (8.2)	35.9 (7.2)	29.0 (5.4)	27.6 (7.6)
Switching Shapes	80.2 (22.0)	60.7 (13.7)	50.9 (11.7)	44.2 (7.3)	41.2 (8.5)
Switching Arrows	75.9 (21.8)	58.5 (11.4)	48.2 (8.3)	40.7 (8.3)	36.9 (8.2)

Table 7. Medians (*Mdn*) and maximums (*Max*) of error scores for the tasks

	7 years	9 years	11 years	13 years	15 years
Task	Mdn (Max)	Mdn (Max)	Mdn (Max)	Mdn (Max)	Mdn (Max)
Naming Shapes	0 (4)	0 (4)	0 (2)	0 (2)	0 (3)
Naming Arrows	0 (5)	0 (2)	0 (2)	0 (2)	0 (2)
Inhibition Shapes	0 (11)	0 (8)	0 (4)	0 (8)	0 (3)
Inhibition Arrows	1 (14)	0 (15)	0 (13)	0 (6)	0 (4)
Switching Shapes	1 (25)	1 (15)	0 (16)	0 (6)	0 (6)
Switching Arrows	2 (15)	1 (9)	1 (12)	0 (8)	0 (5)

Note. The minimum error score was 0 for all groups

The analyses of inhibition scores showed significant differences between the inhibition and switching conditions for both the difference scores, $F(1, 330) = 1005.90$, $p < .001$, $\eta_p^2 = .75$, and the ratio scores, $F(1, 330) = 1204.09$, $p < .001$, $\eta_p^2 = .79$, indicating that the level of difficulty in the switching tasks was larger than in the inhibition tasks. The separate analyses for each inhibition score revealed significant effects of age in all scores: the inhibition difference score, $F(4, 330) = 35.32$, $p < .001$, $\eta_p^2 = .30$; switching inhibition difference score, $F(4, 330) = 63.24$, $p < .001$, $\eta_p^2 = .43$; the inhibition ratio score, $F(4, 330) = 7.06$, $p < .001$, $\eta_p^2 = .08$; and the switching inhibition ratio score, $F(4, 330) = 6.45$, $p < .001$, $\eta_p^2 = .07$ (Figure 2 and Figure 3). The linear and quadratic trends of age were significant in the difference scores, 7-, 9-, and 11-year-olds having significantly larger scores than 15-year-olds. In the ratio scores, the linear trend, but not the quadratic trend, was significant. In the inhibition ratio score, only 7-year-olds had significantly larger ratio scores than 15-year-olds. In the switching inhibition ratio score, 7- and 9-year-olds had significantly larger ratio scores 15-year-olds. The interaction between condition and age was significant in the difference scores, $F(4, 330) = 22.78$, $p < .001$, $\eta_p^2 = .22$, but in follow-up ANOVAs the effect remained significant for all age groups.

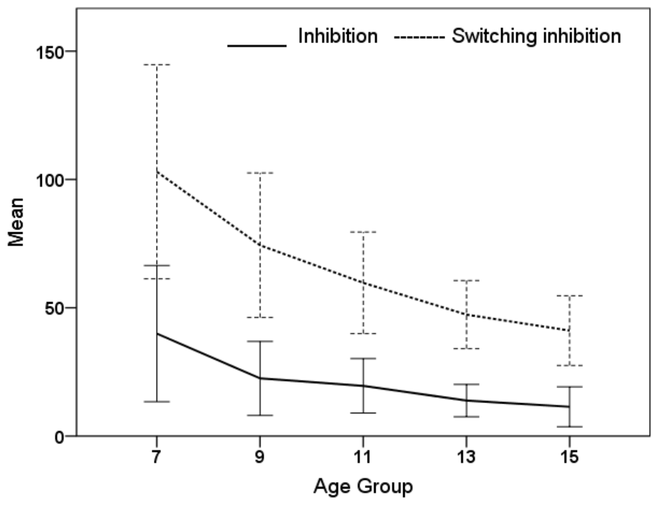

Figure 2. Means and standard deviations (SD +/- 1) of inhibition and switching inhibition difference scores by age group.

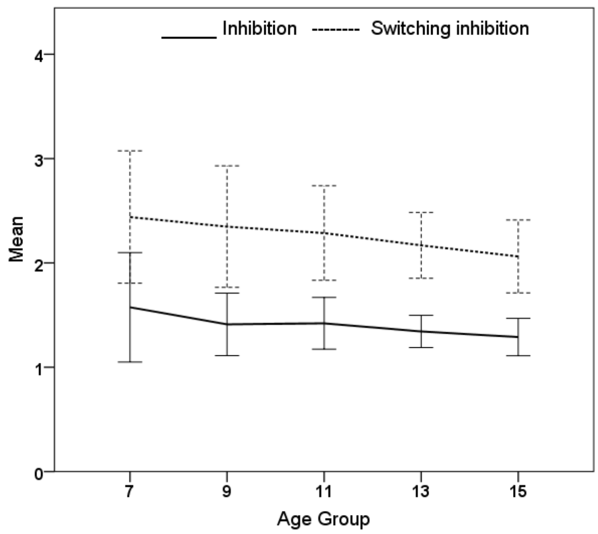

Figure 3. Means and standard deviations (*SD* +/- 1) of inhibition and switching inhibition ratio scores by age group.

In summary, the results of Study II indicate that age-related change in response inhibition continues into the school-age. Age-related improvement in speed continued until age 13, and accuracy improved until age 9 or 13, depending on the complexity of the task. When response inhibition was measured with the difference score, age-related improvement in both inhibition and switching started to decelerate at age 13. With the ratio score, the improvement decelerated earlier and at different ages for different tasks, in inhibition at age 9 and in switching inhibition at age 11. Factors related to the stimuli and task complexity also had effects on response inhibition performance.

4.3 Assessment of executive function behaviors using the ATTEX

The internal consistency of ATTEX total score and each of the ten scales was mostly good both for the normative group, ranging from .73 to .98, and for the combined ADHD group, ranging from .67 to .96.

Effects of sex and age were examined to determine whether separate normative groups for the ATTEX would be required. Sex was significantly associated with the total score, $F(1, 677) = 9.05$, $p = .003$, $\eta_p^2 = .01$, and the scale scores, Wilk's lambda = .92, $F(10, 690) = 5.81$, $p < .001$, $\eta_p^2 = .08$. The follow-up showed that the effect of sex remained significant for each of the scales, boys having higher scores than girls in all scales. Age was significantly associated with only the Motor Hyperactivity scale, $F(9, 687) = 2.80$, $p = .003$, $\eta_p^2 = .04$, 7-year-olds having higher scores than 14-year-olds. Since there were consistent effects of sex but only restricted effects of age on the ATTEX scores, normative data were provided separately for boys and girls, but not for the age groups (Table 8).

Table 8. Normative group mean (*M*) scores, standard deviations (*SD*), and ANOVAs for sex effects

	Total (n=701)		Boys (n=335)		Girls (n=366)				
ATTEX score	M	(SD)	M	(SD)	M	(SD)	F	p	η_p^2
Distractibility	1.1	(1.6)	1.4	(1.7)	0.8	(1.3)	25.6	<.001	.04
Impulsivity	2.1	(3.5)	2.9	(4.0)	1.4	(1.3)	35.9	<.001	.05
Motor Hyperactivity	1.0	(2.1)	1.6	(2.5)	0.5	(1.5)	44.6	<.001	.06
Directing Attention	1.6	(2.2)	2.0	(2.4)	1.3	(2.0)	16.8	<.001	.02
Sustaining Attention	1.3	(2.1)	1.6	(2.3)	1.0	(1.8)	18.3	<.001	.03
Shifting Attention	0.8	(1.5)	1.0	(1.7)	0.6	(1.4)	11.4	.001	.02
Initiative	1.2	(2.0)	1.6	(2.2)	0.9	(1.6)	18.7	<.001	.03
Planning	0.8	(1.5)	1.1	(1.7)	0.6	(1.3)	19.6	<.001	.03
Execution of Action	1.9	(2.8)	2.3	(3.0)	1.5	(2.5)	17.2	<.001	.02
Evaluation	0.6	(1.1)	0.8	(1.3)	0.4	(0.8)	29.9	<.001	.04
Total score	12.7	(17.7)	16.6	(19.8)	9.1	(14.6)	9.1	.003	.01

ANOVA df=1,677 for scales and df=1, 699 for total score

Criterion validity between the ATTEX and the ADHD RS-IV ranged from .58 to .95. The discriminant validity for ADHD diagnoses was examined via two separate ROC analyses. In the first analysis, the estimate of the Area Under the Curve (AUC) was .91 for boys and .93 for girls, indicating a very good concordance between ATTEX total score and ADHD-C/ADHD-I diagnoses (Figure 4). In the second analysis, a diagnostic summary score (ADHD subscore = Directing attention + Planning – Impulsivity – Motor hyperactivity) based on a logistic regression analysis, was used for the ADHD groups. The estimate of AUC was .87, demonstrating good validity for the ADHD subscore in discriminating participants with ADHD-I from participants with ADHD-C.

In addition to the results published in Study III, further analyses were performed to examine the profiles of EF behaviors in the ADHD-C, ADHD-I, and normative groups.

As the groups differed according to the distributions of sex and parent education, these were controlled in the analyses. The effect of age was also controlled because of the large age range in the sample. The ATTEX total scores were significantly higher in the ADHD-C and ADHD-I groups than in the normative group, while the difference between the two ADHD groups was not significant. The overall group difference in the scale scores was also significant, Wilks's lambda = .43, $F(20, 1778) = 46.70$, $p < .001$, $\eta_p^2 = .34$, both ADHD groups having significantly higher scores than the normative group in all scales. Between the ADHD groups, the ADHD-C group had significantly higher scores than the ADHD-I group on two scales, Impulsivity and Motor Hyperactivity, and the ADHD-I group had higher scores on six scales: Directing Attention, Sustaining Attention, Shifting Attention, Initiative, Planning, and Execution of Action (Table 9).

Of the co-occurring disorders, learning disorders were significantly related to the ATTEX total score, $F(1, 208) = 4.44$, $p = .036$, $\eta_p^2 = .02$, children with ADHD plus learning disorders having more (*M* = 64.4, *SD* = 19.7) EF difficulties than other children with ADHD (*M* = 57.6, *SD* = 24.0). The effect was also tested with a model including the group, the non-significant interaction between co-occurring learning disabilities and group indicating that the effect was similar for both ADHD subtypes. In addition, co-occurring cognitive disorders were significantly related to the scale scores, Wilks's lambda = .91, $F(10, 197) = 1.99$, $p = .037$, $\eta_p^2 = .09$, follow-up showing that the association was significant for the Evaluation scale, $F(1, 206) = 4.14$, $p = .043$, $\eta_p^2 = .02$, where children with co-occurring cognitive disorders had more problems (*M* = 3.3, *SD* = 1.8) than other children with ADHD (*M* = 2.7, *SD* = 1.7). Co-occurring disorders of social interaction or conduct disorders were not significantly associated with the ATTEX scores.

In summary, the results of Study III indicated that the ATTEX rating scale has high internal consistency reliability and good criterion and discriminant validity for ADHD. According to the teacher ratings, the ADHD subtypes differed from each other in the EF profiles, and children with predominantly inattentive symptoms showed more wide-ranging difficulties in EFs than children with combined symptoms of inattention and hyperactivity-impulsivity.

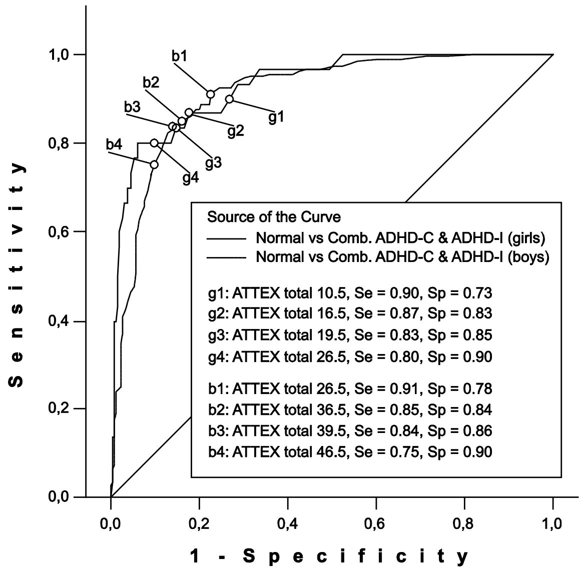

Figure 4. ROC analyses of ATTEX total score for boys (b) and girls (g). Points on the curve indicated by b1-b4 and g1-g4 note the sensitivity (Se) and specificity (Sp) values for selected cutoff scores of ATTEX total scores.

Table 9. ATTEX score means (*M*), standard deviations (*SD*), and ANCOVA results by group

	ADHD-C (C)		ADHD-I (I)		normative (N)					
ATTEX score	*M*	*SD*	*M*	*SD*	*M*	*SD*	$F(2, 898)$	p	η_p^2	Contrasts[a]
Distractibility	5.1	2.0	4.0	2.3	1.1	1.6	337.3	<.001	.43	N < C, I
Impulsivity	11.3	5.1	6.7	4.3	2.1	3.6	302.8	<.001	.40	N < I < C
Motor hyperactivity	6.7	4.1	3.4	3.4	1.1	2.1	251.7	<.001	.36	N < I < C
Directing attention	5.2	2.7	7.8	2.4	1.6	2.2	182.3	<.001	.29	N < C < I
Sustaining attention	5.8	2.9	6.9	3.2	1.3	2.1	255.9	<.001	.36	N < C < I
Shifting attention	4.5	2.4	5.8	2.2	0.8	1.5	320.9	<.001	.42	N < C < I
Initiative	5.0	2.7	7.0	2.6	1.3	2.0	221.9	<.001	.33	N < C < I
Planning	4.0	2.4	4.9	2.3	0.8	1.5	227.5	<.001	.34	N < C < I
Execution of action	7.7	3.9	9.4	3.0	1.9	2.7	246.6	<.001	.36	N < C < I
Evaluation	2.8	1.7	3.2	1.9	0.6	1.1	188.7	<.001	.30	N < C, I
Total score	59.4	23.3	61.4	20.3	12.7	17.7	378.4	<.001	.46	N < C, I

[a] Bonferroni corrected significance level $p < .0167$

5 Discussion

The three studies forming the current thesis addressed assessment and development of EFs in school-age children by using both performance based and behavioral measures. Findings from the developmental studies showed divergent age-related differences for EF tasks of inhibition, attention, and execution of action. The challenges of developmental studies were highlighted in the analyses focusing on response inhibition and showing that both task characteristics and the selection of outcome measures had impact on the developmental findings. A new rating scale that covers a large variety of EF behaviors related to inhibition, attention, and execution of actions was constructed. Investigation of the psychometric properties of the ATTEX demonstrated high internal consistency reliability and good criterion and discriminant validity for ADHD. The following sections include first a more detailed discussion of each study, followed by limitations, general discussion, and practical implications of the results.

5.1 Age-related differences in executive functions

The development of EFs has been extensively studied already for several decades. Like other several other studies in the 1980's and 1990's, Study I examined age-related differences with a wide range of tasks representing different EF domains. The goal in these studies was to investigate the continuum and possible staging of EF development (Becker, Isaac, & Hynd, 1987; Passler et al., 1985; Welsh et al., 1991). The findings from Study I indicated that, when using the subtests from the developmental assessment battery NEPSY, age-related improvement started to decelerate (as compared to the 12-year-old level of performance) first on a task of response inhibition at the age of 6 year. Thereafter, improvement continued on the more complex inhibition task, the planning tasks, and the tasks of selective and sustained attention until ages from 7 to 10 years. In the fluency tasks, the 12-year-old level was reached at the ages between 10 and 11 years. Taking into account the clustering of subtests, the development thus seemed to proceed from the processes of inhibition to attention control and further to fluency.

Previous studies had shown similar differentiation of development in EF task performance, but the developmental sequencing differed across studies (Becker et al., 1987; Passler et al., 1985; Welsh et al., 1991). Various factors related to the study samples, procedures, and measures may account for these differences. For example, differences in sex and parent education level of the samples, which were significantly related to the EF performance on some of the tasks in Study I, could affect the results. Most importantly, however, the developmental findings are influenced by the selection of EF tasks.

In Study I, age-related differences in response inhibition, operationalized as the number of inhibited responses in the Statue subtest, decelerated at the age of 6 years. Becker et al. (1987), using a go-no go task, had shown continuing age-related change in accuracy of inhibition performance until the age of 8 years and in performance speed until the age of 10 years. In a later study, Huizinga et al. (2006) demonstrated how age-related differences in the speed of response inhibition in a Stroop-like task continued until early adulthood. The findings show that the development of response inhibition may appear very different in studies that use different tasks. When using a simple task like the Statue, requiring inhibition of motor responses but involving no response conflict and minimal requirements for working memory, the age-related improvement decelerates early, while performance on a Stroop-like task continues to develop much later. Together, these findings indicate that some relatively simple abilities of response inhibition mature already in the early school years, and the developmental change is much more prolonged for other, more complex capacities related to response inhibition (Garon et al., 2008).

In summary, the findings from Study I were in accordance with previous studies demonstrating that the EFs continue to develop throughout the school-age. The present findings indicated a continuum of development from the processes of response inhibition to attention control, and further to complex EF processes related to effective productivity. Appraisal of the present and other findings on staging of development, however, indicates that the different developmental trends may reflect several other factors than the EFs targeted in the tasks, e.g., degree of difficulty in tasks, the influence of other cognitive processes, or differences in the samples. Thus, it seems important to investigate factors that relate to age-related differences not only *between* the EF domains but also *within* each EF domain.

5.2 Methodological issues affecting the assessment of age-related differences in response inhibition

In Study II, the variation of outcome variables and task characteristics of a Stroop-like response inhibition task highlighted several important methodological issues for the assessment of EFs. The analyses of time and error scores showed that both variables indicated age-related differences in school-age children. The age-related differences in speed were similar in all tasks, performance improving until age 13. The measures for speed thus seemed to reflect age-related differences in the basic naming processes, implicating that outcome variables without control for the effects of non-inhibitory processes are not valid measures of response inhibition.

The error variables have been included only in few previous studies with school age children. In Study II, the 7-year-old group made errors in all tasks, and, in the more complex switching tasks, accuracy continued to increase until age 13. A similar decline of error rates of early school age children was shown also in the recent study of Macdonald et al. (2014) implying that a transition in the automation of response inhibition may take place during this period. Errors also indicate that the tasks are challenging, and an interaction between accuracy and speed may affect the performance (e.g., Davidson et al., 2006). It is thus important that accuracy, alongside with speed, is included as a measure of school age developmental change in inhibition.

The two alternative procedures to control for the effect of non-inhibitory processes indicated different patterns of development. With difference scores, the age-related differences were actually similar to those of the speed measures, while the ratio scores showed earlier decrease in age-related differences at the age of 9 or 11 years. The findings from Study II are in line with the studies of Verhaeghen and colleagues (Schwartz & Verhaeghen, 2008; Verhaeghen & De Meersman, 1998; Verhaeghen & Cerella, 2002), who have demonstrated that the difference score does not eliminate the effect of general processing speed. In clinical adult samples, the relationship between total times for the basic and incongruent naming tasks has been multiplicative, and slower individuals have shown higher difference scores, thus seemingly having more inhibition problems than faster individuals. The present study implies that, in developmental studies with children, a similar effect of basic task speed is involved in the difference scores for inhibition.

The ratio scores have been used only in a few studies on childhood development of response inhibition. In a meta-analysis of studies using the Golden Stroop (Golden, 1978), Schwartz and Verhaeghen (2008) used ratio scores to assess differences between children and adults with ADHD and controls. As an additional result, they found that the ratio scores for reaction times (converted from the original Golden Stroop outcome measures of number of items correctly named in 45 seconds) remained constant over the age range of 9 to 41 years. In the meta-analysis, thus, the ratio score appeared to be immune to age. Similarly, Macdonald et al. (2014) found no age-related differences between ages from 5 to 8 years in the ratio scores of modified pictorial Stroop tasks. Ikeda et al. (2013), however, did show significant age effects in the ratio score for a computerized pictorial Stroop task between a group of 5- to 6-year-olds and older children and also between of 9- to 10-year-olds and adults. These findings indicate that the ratio score is not altogether immune to age. When the effect of processing speed has been controlled for, the developmental change in response inhibition seems to continue at least into the early school years and perhaps even later on.

Task characteristics also had effect on age-related differences. In the tasks involving shapes as stimuli, children mostly performed faster and made fewer errors than in the tasks with arrows. Naming shapes thus seems to be easier and probably more automatized than naming arrows. The distinction between the inhibition scores of the two different Stroop-like tasks implied that the level of difficulty was larger in switching inhibition than in inhibition only. In addition to response inhibition, the switching tasks require shifting of attention (for perception of both the shape/ direction and the color of stimuli), shifting of response set (naming and inhibition), and working memory (two different rules). In the complex switching tasks, the age-related differences cannot be attributed solely to inhibition, as they may actually reflect differences in these other cognitive processes. Factor related to the tasks stimuli, rules, and requirements should be carefully documented and taken into account when integrating findings across developmental studies.

In summary, several steps are necessary when choosing the measures for developmental studies of response inhibition. First, age-related differences in both speed and accuracy should be examined, and the effect of errors should be included in the calculated inhibition scores. Second, the effects of other cognitive processes than inhibition should be controlled for. A minimum requirement is the appliance of difference scores that control for the basic naming process, but do not eliminate the effect of general processing speed (Lansbergen et al., 2007). The ratio scores, or other proportional scores, can be used to minimize the processing speed effects. Finally, awareness of other factors that may contribute to performance on inhibition tasks, e.g., ceiling effects, is required.

5.3 The ATTEX and behavioral executive function difficulties in ADHD

Assessment of behavioral EFs is an essential part of clinical assessment of ADHD and many other developmental disorders. Study III examined the properties of a new rating scale, the ATTEX, for assessment of EFs in school environments.

The study provided normative data for Finnish 7- to 15-year-old children. Similarly with other normative studies (DuPaul et al., 1998; Gioia et al., 2000), boys had higher scores than girls in the ATTEX ratings. To avoid sex-related bias in ratings, both combined norms and separate norms for boys and girls were provided for ATTEX. Using norms not differentiated for sex would present a risk for under-identifying girls with EF difficulties. On the other hand, using sex-specific norms alone would result in an equal percentage of ADHD-diagnosed boys and girls, thus risking an over-identification of girls with ADHD (Collett et al., 2003). Access to both sex-specific and combined norms creates an ideal opportunity for evaluating the severity of a child's symptoms relative to the population of peers while also placing their symptoms in the sex-specific context.

Somewhat unexpectedly, the effect of age was significant only for the Motor Hyperactivity scale and only between two age groups, 7-year-olds having more problems of hyperactivity than 14-year-olds. Previous studies have indicated that younger children have higher ratings on ADHD and EF scales than older children and adolescents (DuPaul et al., 1998; Gioia et al., 2000). Although restricted to two age groups, results of the present study showed a similar tendency. They were also in line with studies showing that motor hyperactivity decreases with the increase of age (Larsson, Lichtenstein, & Larsson, 2006; Seidman, 2006).

The investigation of psychometric properties of ATTEX showed high internal consistency reliability and good criterion validity. The discriminative potential for clinical groups of ADHD was also good, the summary total score showing validity for discriminating children with ADHD from normal controls, and the ADHD subscore discriminated children with ADHD-C from children with ADHD-I. The ATTEX can thus be used as an accurate tool in screening for attention disorders and in differentiating ADHD-I from ADHD-C. On the bases of ROC analyses, suitable cutoff scores can be selected for different purposes. For screening to rule out the possibility of ADHD, a cutoff score with high specificity (the probability that a child who does not have ADHD is rated below the cutoff score) would be optimal. Accordingly, for diagnostic purposes or ruling in ADHD, a cutoff score with high sensitivity would best serve the purposes.

For clinical purposes, e.g., planning for interventions, it is important that the rating scale gives detailed information of the EF behaviors. Closer examination of ATTEX scales revealed that the EF profiles of children with ADHD-C and ADHD-I were different. For children with ADHD-C, the teacher ratings of EFs reflected difficulties related to problems in hyperactivity and impulsivity. Contrary to previous findings (Gioia et al., 2002; McCandless & O' Laughlin, 2007; Semrud-Clikeman et al., 2010), the EF difficulties of children with ADHD-I appeared more wide-ranging. The ATTEX ratings indicated difficulties in attention regulation as well as in taking initiative, planning, and execution of action. Although not indicated in previous studies using EF ratings, the wide-ranging difficulties may relate to slow motor output and cognitive tempo of children with ADHD-I (Carlson & Mann, 2002; Nigg, Blaskey, Huang-Pollock, & Rappley, 2002). The EF difficulties of the ADHD-I group may also come out more clearly in school environments that set specific requirements for sustained attention and self-directed initiative and execution of action. Accordingly, teachers may have more opportunities to observe these behaviors than parents do.

In summary, the ATTEX showed good internal consistency and good validity in discriminating children with ADHD from non-ADHD children and also in differentiating between the ADHD subtypes. The profiles of EF behaviors indicated that children with predominantly inattentive symptoms of ADHD may have even more wide ranging EF impairments than children with combined symptoms in school situations. Future studies examining the test-retest reliability and confirmatory factor analysis are needed for to evaluate the stability of the measure the internal structure of ATTEX. The findings concerning EF profiles in ADHD subtypes need to be replicated with other samples, preferably including larger ADHD-I groups.

5.4 Limitations of the studies

It is important to keep in mind that the samples in Studies I and II, like in most other developmental studies on EFs, were cross-sectional. The effects of individual differences among the participants within each age group can be fully controlled for only in longitudinal studies. Additionally, in the sample of Study II, mother’s education level centered on medium level, indicating less variability than would be found in the population. As parent education level was related to EF performance in Study I and also in other previous studies (Ardila, Rosselli, Matute, & Guajardo, 2005), the results of Study II may not fully reflect the diversity in the development of response inhibition.

In Study III, characteristics of the clinical sample may limit generalization of the presented EF profiles. In accordance with earlier findings (Mattison & Mayes, 2012; Rucklidge & Tannock, 2002), the analyses of co-occurring disorders showed that learning disorders had an additive effect on the EF difficulties of children with ADHD. As the majority of children in the ADHD-I group had co-occurring learning disorders, the wide-ranging EFs could reflect this effect. Because of the small group size, however, the groups with and without learning disorders could not be compared within the ADHD-I group. Overall, the ADHD-I group was notably smaller than would be expected according to prevalence studies (Froehlich et al., 2007; Skounti, Philalithis, & Galanakis, 2007). It is reasonable to assume that children with severe symptoms of inattention and complicated comorbidities of learning disabilities would have been referred to the child neurology special unit. Thus, the wide-ranging difficulties in EF behaviors may be typical only for the most severe cases of ADHD-I. Further, because of the small size of the ADHD-I group, statistical tests may underestimate the EF deficits in the ADHD-I group versus the other two groups.

Limitations related to the validity of EF tasks also need to be considered when interpreting the findings. In Study I, the effect of other factors than the intended EFs were not controlled for. In addition to the effect of other cognitive processes, the validity of tasks may depend on the age of the participants. The task sensitivity to age-related change has a fundamental effect on developmental findings. In Study II, using manual response inhibition tasks may present a ceiling effect for the oldest age groups. Specifically, computerized tasks may be more sensitive to subtle age-related differences as they measure response time and latency very precisely and allow analyses that dissociate between the different stages of the response (e.g., Bub, Masson, & Lalonde, 2006).

As for the EF ratings, it is important to keep in mind that source effects related to the rater may reflect on the results. In Study III, the teachers may have been influenced by knowing that the child is on medication or has been evaluated for attention disorders.

As compared to the later studies, the data analyses in Study I were also limited in that they did not include measures of effect size.

5.5 General discussion and practical implications

Stretching over a period of more than ten years, this thesis highlights some aspects of the recent advances in EF measurement. Unlike former studies on normative EF development that relied on mixed batteries of EF tasks (e.g., Becker et al., 1987; Passler et al., 1985; Welsh et al., 1991), Study I employed the EF and attention tasks from the NEPSY, first published in 1997. Derived from Luria's theory and prevailing traditions of child neuropsychological assessment, the tasks comprised a comprehensive and consistent assessment of EFs. The tasks were modified for preschool- and school-age children and were intended to be fairly simple tasks focusing on certain EFs. However, some of them, e.g., the Tower, were still relatively complex EF tasks. In NEPSY-II, published in 2007, the measures were further refined by adding items to enhance sensitivity to older and younger age groups, by omitting tasks with less clinical sensitivity, and by adding new tasks to improve the coverage of EFs (Korkman, Kirk, & Kemp, 2007b). An important improvement in the newer EF test batteries (e.g., the Delis-Kaplan Executive Function System D-KEFS; Delis et al., 2001) has been the inclusion of contrast scores that control for the effects of non-executive processes on test performance. In NEPSY-II, the new Inhibition subtest was the first task that included a basic task that could be used as a control task. This allowed a more precise measurement of the targeted EFs and increased the validity of measurement.

Study II employed two different methods of controlling for the effects of non-inhibitory processes (the US NEPSY-II involves another controlling method not included in these analyses, for more information see Korkman et al., 2007b). Exploration of the different outcome variables showed how important these controlling measures are for assessment of inhibition. For clinical assessment of inhibition, the findings imply that the performance on the individual inhibition tasks is not a reliable measure of response inhibition and controlling measures should be used when interpreting the results. Especially, if the child's performance in the basic naming task is slow, the ratio score should be used to control for the effects of general slowness.

The behavioral assessment of EFs in Study III represented another aspect of advancement in the validity of EF measures. Behavioral rating scales have provided an alternative method for the assessment of EFs with more ecological validity for predicting children's EF difficulties in the everyday environment. In clinical assessment of EFs, tests may provide relevant information of some EF domains, e.g., working memory, and they are useful in the assessment of difficulties and strengths in other cognitive functions. For the evaluation of how EFs affect the child's functioning, ratings from different environments should be included in the assessment procedure. The new rating scale ATTEX was purported specifically for assessment of EFs in school environments. In health-care units, the ATTEX can be used as part of the diagnostic procedure for ADHD. It can also be used by school psychologist in screening for EF difficulties. The results from Study III implicate that the wide-ranging ratings may be sensitive to individual differences and thus may provide a comprehensive description of EF behaviors. As the teacher-completed ratings give direct information of the classroom teacher's observations of the student's difficulties, the ATTEX may be useful in the collaborative work among school professionals (DuPaul, Weyandt, & Janusis, 2011). On the basis of the structured and detailed EF profiles, teachers and school psychologists can work in collaboration when planning and implementing school based interventions.

The findings from Studies I and II can also be viewed in relation to the theoretical models of EFs. In Study I, the exploratory factor structure and the developmental sequencing implicated relative separateness of performance in tasks of response inhibition, attention, and more complex processes of fluency. This is in line with Barkley's (1997) model of EFs and developmental studies that emphasize the primary nature of inhibition and attention as core functions of EFs. However, findings based on exploratory factor analysis are highly dependent on the measures that are employed, and the low correlations between measures, interpreted as separateness of functions, may also arise from differences in non-executive processing requirements or low reliabilities of the measures (Miyake et al., 2000). As has been found in both exploratory and confirmatory analyses of EF performance, the obtained factors vary across studies (Huizinga et al., 2006; Levin et al., 1991; Taylor et al., 1996; Welsh et al., 1991). Further, when using data from several different age groups, the factors obtained in this study - at least to some extent - reflect the influence of age. More recent studies have clearly indicated that correlations among EF task performances change during development, and the factor structure is different for different age groups (e.g., Lee et al., 2013). In Study I, the factor structure for different age groups could not be examined because of small sample sizes.

The findings from Study II highlight the role of processing speed in inhibition performance. Processing speed is often seen as having a confounding effect on inhibition. In information processing models, however, speed can be seen as including in the inhibitory processes. For example, in the model of Friedman and Miyake (2004), inhibition proceeds in stages, starting from inhibition of external distractors in the selection of relevant information, proceeding to inhibition of irrelevant thoughts and ideas to protect the working memory, and finally proceeding to inhibition of incorrect responses that enables performance of relevant responses. It would seem logical that speed of inhibition is an important factor in each stage of inhibition. For example in response inhibition, speed in the selection of relevant motor output could be essential in the constantly changing real-world situations. For research on development of inhibition, the implications are twofold. In developmental studies, the difference score has been the standard procedure in controlling for the effects of basic processes. If the effect of processing speed is seen as confounding, then the outline of developmental change in response inhibition may need re-evaluation. If, on the other hand, response inhibition is seen as including the effect of processing speed, then the findings that reflect developmental change may be appropriate. The findings from Study II indicate that the choice of an appropriate control method may depend on the focus of the study. If groups that are very different in the basic processing speed are compared, e.g., ADHD *vs.* controls, then ratio scores should be used. If the purpose is to gain knowledge of developmental change, then it may be reasonable to apply also the difference score that reflects the change in the speed of inhibition.

5.6 Conclusions

The cognitive processes and behaviors related to EFs develop throughout childhood. In line with previous studies, this thesis showed that the developmental proceeding varies across the different EF processes. The relative differences between EF domains, however, may actually reflect the characteristics of the measures more than the EF constructs as such. The closer examination of developmental variation in response inhibition demonstrated that the development of EFs is more accurately portrayed as proceeding from simple forms of processing in straightforward contexts to the more complex manifestations of the function in more complicated contexts. For example, simple inhibition of motor responses (e.g., keeping seated for a required period of time) can be mastered already in preschool age. The ability to both withhold responses and execute other responses according to instructions (e.g., keeping seated while listening, withholding any spontaneous questions or comments, and acting upon the instructions) continues to develop into the later school age. The sequence of development within each EF domain may thus be more important than the possible staging between the different domains. In addition to the actual processes of EFs, several factors related to other cognitive processes as well as task materials, stimuli, and the selected outcome measures have effect on how the development is depicted. In future studies, careful examination of these factors can help to attain a more consisted account on EF development.

Knowledge of the factors that affect the validity of measures is equally important for clinicians. Many of the measures that are employed in developmental studies are also employed in the clinical assessment of EF difficulties. In addition, the clinical measures need to be sensitive to the actual difficulties that arise in every-day situations. These are best measured with standardized rating scales. The new rating scale presented in this thesis provides normative data for Finnish school-age children and has appropriate psychometric properties for assessment of EF behaviors in school environments. In the clinical assessment of EFs, information from different environments and different informants is needed to obtain a comprehensive evaluation. Standardized rating scales for parents and youth's self-evaluation as well for younger preschool-age children and also for adults are still needed.

6 References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders DSM-IV*. Washington (D.C.): American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders DSM-5*. Washington (D.C.): American Psychiatric Publishing.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology, 8*(4), 231–240.

Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology, 28*(1), 539–560. doi:10.1207/s15326942dn2801_5

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49A*(1), 5–28.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65–94.

Barkley, R. A. (2006). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York: The Guilford Press.

Barkley, R. A. (2012). *Barkley deficits in executive function scale-children and adolescents*. New York: Guilford Press

Becker, M. G., Isaac, W., & Hynd, G. W. (1987). Neuropsychological development of nonverbal behaviors attributed to "frontal lobe" functioning. *Developmental Neuropsychology, 3*(3-4), 275–298.

Ben-David, B. M., Nguyen, L. L. T., & van Lieshout, P. H. H. M. (2011). Stroop effects in persons with traumatic brain injury: Selective attention, speed of processing, or color-naming? A meta-analysis. *Journal of the International Neuropsychological Society, 17*(2), 354–363.

Bennett, P. C., Ong, B., & Ponsford, J. (2005). Assessment of executive dysfunction following traumatic brain injury: Comparison of BADS with other clinical neuropsychological measures. *Journal of the International Neuropsychological Society, 11*, 606–613.

Benton, A. L., & Hamsher, K. d. (1989). *Multilingual aphasia examination*. Iowa City: AJA Associates.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641–1660. doi:10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning & Individual Differences, 21*(4), 327–336. doi:10.1016/j.lindif.2011.01.007

Biederman, J. (2008). Discordance between psychometric testing and questionnaire-based definitions of executive function deficits in individuals with ADHD. *Journal of Attention Disorders, 12*(1), 92–102.

Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., et al. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting & Clinical Psychology, 72*(5), 757–766.

Biederman, J., Petty, C., Fried, R., Fontanella, J., Doyle, A. E., Seidman, L. J., & Faraone, S. V. (2006). Impact of psychometrically defined deficits of executive functioning in adults with attention deficit hyperactivity disorder. *American Journal of Psychiatry, 163*(10), 1730–1738.

Booth, J. N., Boyle, J. M. E., & Kelly, S. W. (2010). Do tasks make a difference? accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *British Journal of Developmental Psychology, 28*(1), 133–176. doi:10.1348/026151009X485432

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. doi:10.1016/0022-0965(82)90008-X

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence, 47*, 34–43. doi: 10.1016/j.intell.2014.08.010

Bub, D. N., Masson, M. E. J., & Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychological Science, 17*(4), 351–357. doi:10.1111/j.1467-9280.2006.01710.x

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.
Carlson, C. L., & Mann, M. (2002). Sluggish cognitive tempo predicts a different pattern of impairment in the attention deficit hyperactivity disorder, predominantly inattentive type. *Journal of Clinical Child & Adolescent Psychology, 31*(1), 123–129.
Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595–616. doi:10.1207/s15326942dn2802_3
Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. doi:10.1111/1467-8624.00333
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed. ed.). Hillsdale, NJ: Erlbaum.
Collett, B. R., Ohan, J. L., & Myers, K. M. (2003). Ten-year review of rating scales. V: Scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(9), 1015–1037.
Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology, 31*(1), 103–128. doi:10.1207/s15326942dn3101_6
Conners, C. K. (1995). *The Conners continuous performance test*. Toronto: Multi-Health Systems.
Davidson, M., Amso, D., Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037–2078.
Delis, D. C. (2012). *Delis rating of executive functions*. Bloomington, MN: Pearson.
Delis, D. C., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan Executive Function System*. San Antonio, TX: The Psychological Corporation.
Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development, 56*(4), 868–883. doi:10.1111/1467-8624.ep7251346
Dimoska-Di Marco, A., McDonald, S., Kelly, M., & Tate,R., Johnstone,S. (2011). A meta-analysis of response inhibition and stroop interference control deficits in adults with traumatic brain injury (TBI). *Journal of Clinical and Experimental Neuropsychology, 33*(4), 471–485.
Doyle, A. E., Faraone, S. V., Seidman, L. J., Willcutt, E. G., Nigg, J. T., Waldman, I. D., et al. (2005). Are endophenotypes based on measures of executive functions useful for molecular genetic studies of ADHD? *Journal of Child Psychology & Psychiatry & Allied Disciplines, 46*(7), 774–803.
DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD rating scale -IV: Checklists, norms, and clinical interpretation*. New York: The Guilford Press.
DuPaul, G. J. (2003). Assessment of ADHD symptoms: Comment on gomez et al. (2003). *Psychological Assessment, 15*(1), 115–117.
DuPaul, G. J., Weyandt, L. L., & Janusis, G. M. (2011). ADHD in the classroom: Effective intervention strategies. *Theory into Practice, 50*(1), 35–42. doi:10.1080/00405841.2011.534935
Emslie, H., Wilson, F. C., Burden, V., Ninno-Smith, I., & Wilson, B. A. (2003). *Behavioural assessment of the dysexecutive syndrome for children (BADS-C)*. London, UK: Harcourt Assessment.
Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101–135.
Froehlich, T. E., Lanphear, B. P., Epstein, J. N., Barbaresi, W. J., Katusic, S. K., & Kahn, R. S. (2007). Prevalence, recognition, and treatment of attention-Deficit/Hyperactivity disorder in a national sample of US children. *Archives of Pediatrics and Adolescent Medicine, 161*(9), 857–864.
Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60.
Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 312–7 years old on a stroop- like day-night test. *Cognition, 53*(2), 129–153. doi:10.1016/0010-0277(94)90068-X
Gioia, G., Isquith, P., Guy, S., & Kenworthy, L. (2000). *BRIE The behavior rating inventory of executive function*. Lutz, FL: Psychological Assessment Resources.
Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of everyday executive function in acquired and development disorders. *Child Neuropsychology, 8*(2), 121–137.
Golden, C. J. (1978). *Stroop color and word test: A manual for clinical and experimental uses*. Chicago: Stoelting.
Goldman-Rakic, P. (1987). Development of cortical circuitry and cognitive function. *Child Development, 58*(3), 601–622. doi:10.1111/1467-8624.ep7264449

Gomez, R. (2007). Australian parent and teacher ratings of the DSM-IV ADHD symptoms: Differential symptom functioning and parent-teacher agreement and differences. *Journal of Attention Disorders, 11*(1), 17–27.

Gomez, R., Leonard Burns, G., Walsh, J. A., & Alves de Moura, M. (2003). A multitrait-multisource confirmatory factor analytic approach to the construct validity of ADHD rating scales. *Psychological Assessment, 15*(1), 3–16. Retrieved from

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586. doi:10.1111/j.1469-7610.1997.tb01545.x

Grant, D. A., & Berg, E. (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, 38*(4), 404–411.

Halperin, J. M.,Trampush, J. W., Miller, C. J., Marks, D. J., & Newcorn, J. H. (2008). Neuropsychological outcome in adolescents/young adults with childhood ADHD: Profiles of persisters, remitters and controls. *Journal of Child Psychology and Psychiatry, 49*(9), 958–966.

Herba, C. M., Tranah, T., Rubia, K., & Yule, W. (2006). Conduct problems in adolescence: Three domains of inhibition and effect of gender. *Developmental Neuropsychology, 30* , 659–695.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences, 8*(1), 26–32.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*(2), 233–253. doi:10.1111/j.2044-835X.1998.tb00921.x

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*(11), 2017–2036.

Hummer, T. A. (2011). Executive functioning characteristics associated with ADHD comorbidity in adolescents with disruptive behavior disorders. *Journal of Abnormal Child Psychology, 39*(1), 11–19. doi:10.1007/s10802-010-9449-3

Ikeda, Y., Okuzumi, H., & Kokubun, M. (2013). Age-related trends of stroop-like interference in animal size tests in 5- to 12-year-old children and young adults. *Child Neuropsychology, 19*(3), 276–291. doi:10.1080/09297049.2012.658364

Jarratt, K. P., Riccio, C. A., & Siekierski, B. M. (2005). Assessment of attention deficit hyperactivity disorder (ADHD) using the BASC and BRIEF. *Applied Neuropsychology, 12*(2), 83–93. doi:10.1207/s15324826an1202_4

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review, 17*(3), 213–233.

Kadesjö, B., Janols, L., Korkman, M., Mickelson, K., Strand, G., & Trillingsgaard, A. (2004). The FTF (five to fifteen): The development of a parent questionnaire for the assessment of ADHD and comorbid conditions. *European Child and Adolescent Psychiatry, 13*(Suppl. 3), 3–13.

Keith, R. W. (1994). *The auditory continuous performance test.* San Antonio, TX: Psychological Corporation.

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

Klenberg, L., Jämsä, S., Häyrinen, T., Lahti-Nuuttila, P., & Korkman, M. (2010). The attention and executive function rating inventory (ATTEX): Psychometric properties and clinical utility in diagnosing ADHD subtypes. *Scandinavian Journal of Psychology, 51,* 439–448.

Klenberg, L., Jämsä, S., Häyrinen, T., & Korkman, M. (2010). *Keskittymiskysely*. Helsinki: Psykologien Kustannus.

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.

Kochanska, G., Tjebkes, J. L., & Forman, D. R. (1998). Children's emerging regulation of conduct: Restraint, compliance, and internalization from infancy to the second year. *Child Development, 69*(5), 1378-1389. doi:10.2307/1132272

Korkman, M., Kirk, U., & Kemp, S. (1997). *NEPSY. Lasten neuropsykologinen tutkimus.* Helsinki: Psykologien Kustannus.

Korkman, M., Kirk, U., & Kemp, S. L. (1998). *NEPSY. A developmental neuropsychological assessment.* San Antonio, TX: Psychological Corporation.

Korkman, M., Jaakkola, M., Ahlroth, A., Pesonen, A. E., & Turunen, M. M. (2004). Screening of developmental disorders in five-year-olds using the FTF (five to fifteen) questionnaire: A validation study. *European Child & Adolescent Psychiatry, 13*(Suppl. 1.), 31–38.

Korkman, M., Kirk, U., & Kemp, S. L. (2007a). *NEPSY II. Administrative manual.* San Antonio, TX: PsychCorp. A Brand of Harcourt Assessments.
Korkman, M., Kirk, U., & Kemp, S. L. (2007b). *NEPSY II. Clinical and interpretative manual.* San Antonio, TX: PsychCorp. A Brand of Harcourt Assessments.
Korkman, M., Kirk, U., & Kemp, S. L. (2008). *NEPSY II - Lasten neuropsykologinen tutkimus.* Helsinki: Psykologien Kustannus Oy.
Koskelainen, M., Sourander, A., & Kaljonen, A. (2000). The strengths and difficulties questionnaire among finnish school-aged children and adolescents. *European Child & Adolescent Psychiatry, 9*(4), 277–284.
Lansbergen, M. M., Kenemans, J. L., & van Engeland, H. (2007). Stroop interference and attention-Deficit/Hyperactivity disorder: A review and meta-analysis. *Neuropsychology, 21*(2), 251-262.
Larsson, H., Lichtenstein, P., & Larsson, J. O. (2006). Genetic contributions to the development of ADHD subtypes from childhood to adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(8), 973–981.
Lawrence, V., Houghton, S., Douglas, G., Durkin, K., Whiting, K., & Tannock, R. (2004). Executive function and ADHD: A comparison of children's performance during neuropsychological testing and real-world activities. *Journal of Attention Disorders, 7*(3), 137–149.
Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development, 84,* 1933–195. doi: 10.1111/cdev.12096
Leon-Carrion, J., García-Orza, J., & Perez-Santamaría, F. J. (2004). Development of the inhibitory component of the executive functions in children and adolescents. *International Journal of Neuroscience, 114*(10), 1291–1311. doi:10.1080/00207450490476066
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. doi:10.1080/87565649609540638
Lezak, M. D. (1993). Newer contributions to the neuropsychological assessment of executive functions. *Journal of Head Trauma Rehabilitation, 8*, 24–31.
Lezak, M. D. (1995). *Neuropsychological assessment*. New York: Oxford University Press.
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.
Luria, A. R. (1980). *Higher cortical functions in man*. New York: Basic Books.
Luria, A. R. (1973). *The working brain. An introduction to neuropsychology*. New York: Basic Books.
Macdonald, J. A., Beauchamp, M. H., Crigan, J. A., & Anderson, P. J. (2014). Age-related differences in inhibitory control in the early school years. *Child Neuropsychology, 20*, 509–526. doi:10.1080/09297049.2013.822060
Mares, D., McLuckie, A., Schwartz, M., & Saini, M. (2007). Executive function impairments in children with attention-deficit hyperactivity disorder: Do they differ between school and home environments? *Canadian Journal of Psychiatry, 52*(8), 527–534.
Mattison, R. E., & Mayes, S. D. (2012). Relationships between learning disability, executive function, and psychopathology in children with ADHD. *Journal of Attention Disorders, 16*(2), 138–146. doi:10.1177/1087054710380188
McAuley, T., Christ, S. E., & White, D. A. (2011). Mapping the development of response inhibition in young children using a modified day-night task. *Developmental Neuropsychology, 36*(45), 539–551.
McAuley, T., & White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology, 108*(3), 453–468. doi:10.1016/j.jecp.2010.08.009
McCandless, S., & O' Laughlin, L. (2007). The clinical utility of the behavior rating inventory of executive function (BRIEF) in the diagnosis of ADHD. *Journal of Attention Disorders, 10*(4), 381–389.
Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B.,& Kellam, S. G. (1991). Analysis of elements of attention: a neuropsychological approach. *Neuropsychology Review, 2,*109–145.
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.
Myers, K., & Winters, N. C. (2002). Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(2), 114–122. doi:10.1097/00004583-200202000-00004

Naglieri, J. A., & Goldstein, S. (2013). *Comprehensive executive functioning index*. Toronto: Multi Health Systems.

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. doi:10.1037/0033-2909.126.2.220

Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and DSM-IV ADHD subtypes. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(1), 59–66.

Nigg, J. T., & Casey, B. J. (2005). An integrative theory of attention-deficit/ hyperactivity disorder based on the cognitive and affective neurosciences. *Development & Psychopathology, 17*(3), 785–806.

Nigg, J. T. (2001). Is ADHD as disinhibitory disorder? *Psychological Bulletin, 127*(5), 571–598.

Official Statistics of Finland. (2012). *Educational structure of population [e-publication]. ISSN=2242-2919. helsinki: Statistics finland.* Retrieved 09/17, 2013, from http://tilastokeskus.fi/til/vkour/meta_en.html.

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 1*(4), 349–370.

Posner, M. I., & Rothbart, M. K. (1998). Attention, self–regulation and consciousness. *Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences, 353*(1377), 1915–1927. doi:10.1098/rstb.1998.0344

Posner, M. I., Sheese, B. E., Odludaş, Y., & Tang, Y. (2006). Analyzing and shaping human attentional networks. *Neural Networks, 19*(9), 1422–1429. doi:10.1016/j.neunet.2006.08.004

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of Experimental Child Psychology, 108*(3), 621–637.

Rebok, G. W., Smith, C. B., Pascualvaca, D. M., Mirsky, A. F., Anthony, B. J., & Kellam, S. G. (1997). Developmental changes in attentional performance in urban children from eight to thirteen years. *Child Neuropsychology, 3*(1), 28–46.

Regard, M., Srauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual & Motor Skills, 55*(3), 839–844.

Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied Neuropsychology, 12*(4), 190–201.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*, 343–350.

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry, 43*, 988–1003. doi:10.1111/1469-7610.00227

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573–594. doi:10.1207/s15326942dn2802_2

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology, 39*(5), 877–890.

Schwartz, K., & Verhaeghen, P. (2008). ADHD and stroop interference from age 9 to age 41 years: A meta-analysis of developmental effects. *Psychological Medicine, 38*(11), 1607–1616.

Seidman, L. J. (2006). Neuropsychological functioning in people with ADHD across the lifespan. *Clinical Psychology Review, 26*(4), 466–485.

Semrud-Clikeman, M., Walkowiak, J., Wilkinson, A., & Butcher, B. (2010). Executive functioning in children with asperger syndrome, ADHD-combined type, ADHD-predominately inattentive type, and controls. *Journal of Autism and Developmental Disorders, 40*(8), 1017–1027.

Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-Deficit/Hyperactivity disorder? *Behavioural Brain Research, 130*(1–2), 3–28. doi:10.1016/S0166-4328(01)00430-2

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London.B, Biological Sciences, 298*(1089), 199–209. doi:10.1098/rstb.1982.0082

Skounti, M., Philalithis, A., & Galanakis, E. (2007). Variations in prevalence of attention deficit hyperactivity disorder worldwide. *European Journal of Pediatrics, 166*(2), 117–123. doi:10.1007/s00431-006-0299-5

Sonuga-Barke, E. J. S. (2005). Causal models of attention-Deficit/Hyperactivity disorder: From common simple deficits to multiple developmental pathways. *Biological Psychiatry, 57*(11), 1231–1238.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662.
St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, *59*, 745–759.
Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.
Sullivan, J. R., & Riccio, C. A. (2007). Diagnostic group differences in parent and teacher ratings on the BRIEF and conners' scales. *Journal of Attention Disorders*, *11*(3), 398–406.
Thorell, L. B., & Nyberg, L. (2008). The childhood executive functioning inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology*, *33*(4), 536–552.
Toll, S. W. M., Van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, *44*(6), 521–532. doi:10.1177/0022219410387302
Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2009). Executive functions: Performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*(1), 53–72.
van Mourik, R., Oosterlaan, J., & Sergeant, J. A. (2005). The stroop revisited: A meta-analysis of interference control in AD/HD. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *46*(2), 150–165.
van Zomeren, A. H., & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*. New York: Oxford University Press.
Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience & Biobehavioral Reviews*, *26*(7), 849–57.
Verhaeghen, P., & De Meersman, L. (1998). Aging and the stroop effect: A meta-analysis. *Psychology and Aging*, *13*(1), 120–126. doi:10.1037/0882-7974.13.1.120
Vik, P., & Ruff, R. R. (1988). Children's figural fluency performance: Development of strategy use. *Developmental Neuropsychology*, *4*(1), 63–74.
Watson, A. J., & Bell, M. A. (2013). Individual differences in inhibitory control skills at three years of age. *Developmental Neuropsychology*, *38*(1), 1–21. doi:10.1080/87565641.2012.718818
Wechsler, D. (1955). *The wechsler adult intelligence scale.* New York: The Psychological Corporation.
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.
Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, *108*, 436–452. doi:10.1016/j.jecp.2010.08.008
Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-Deficit/Hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*(11), 1336–1346.
Wu, K. K., Chan, S. K., Leung, P. W. L., Liu, W., Leung, F. L. T., & Ng, R. (2011). Components and developmental differences of executive functioning for school-aged children. *Developmental Neuropsychology*, *36*(3), 319–337. doi:10.1080/87565641.2010.549979