

**Cognitive and emotional development in children and adolescents with  
ADHD compared to typical developing children.**

**Erik Winther Skogli**

Innlandet Hospital Trust Lillehammer, Division Mental Health Care

Anders Sandvigsgate 17, 2629 Lillehammer

and

Department of Psychology

University of Oslo

Lillehammer, 2014

© Erik Winther Skogli, 2014

*Series of dissertations submitted to the  
Faculty of Social Sciences, University of Oslo  
No. 474*

ISSN 1504-3991

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.  
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing, Oslo.  
The thesis is produced by Akademika Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.



## CONTENTS

<b>Acknowledgements</b> .....	4
<b>Summary</b> .....	5
<b>List of papers</b> .....	7
<b>Abbreviations</b> .....	8
<b>1. INTRODUCTION</b> .....	9
1.1 ADHD and diagnostic criteria .....	10
1.2 Comorbidity in ADHD .....	11
1.3 Gender differences in ADHD .....	11
1.4 Executive dysfunction – a core feature in children and adolescents with ADHD. ....	12
1.4.1 Barkley’s model of a core deficit in response inhibition (1997).....	12
1.4.2 Sonuga-Barke’s dual pathway model of ADHD (2003).....	13
1.4.3 Cold executive function in children and adolescents with ADHD .....	14
1.4.4 Hot executive function in children and adolescents with ADHD .....	14
1.4.5 Executive function in everyday situations in children and adolescents with ADHD .....	15
1.5 Typical development of executive function .....	16
1.5.1 Development of cold executive function in children and adolescents with ADHD .....	16
1.5.2 Development of hot executive function in children and adolescents with ADHD.....	17
1.6 Unanswered questions.....	17
<b>2. AIMS</b> .....	18
<b>3. METHODS</b> .....	18
3.1 Design .....	18
3.2 Procedure .....	19
3.3 Participants .....	19
<b>3.4 Measures</b> .....	20
3.4.1 Clinical assessment .....	20
3.4.2 Neurocognitive assessment .....	21
3.4.3 Parent-rated executive function.....	23
3.4.4 Symptom assessment .....	24

3.4.4.1	Parent-rating scales .....	24
3.4.4.2	Self-report scales .....	24
<b>3.5</b>	<b>Statistical analyses</b> .....	25
<b>4.</b>	<b>SUMMARY OF PAPERS</b> .....	27
<b>5.</b>	<b>DISCUSSION</b> .....	29
<b>5.1</b>	<b>Main findings</b> .....	30
5.1.1	Hot and cold executive function in children and adolescents with ADHDsubtypes.....	30
5.1.2	Executive dysfunction and comorbidity in girls and boys with ADHD ....	34
5.1.3	Development of hot and cold executive function in boys and girls with ADHD.....	35
<b>5.2</b>	<b>Methodological issues</b> .....	38
5.2.1	Representativeness and generalisability of results .....	38
5.2.2	Possible confounding factors .....	38
5.2.3	Attrition.....	40
5.2.4	Measurement.....	40
<b>5.3</b>	<b>Implications</b> .....	42
5.3.1	Theoretical implications .....	42
5.3.2	Clinical implications .....	42
<b>5.4</b>	<b>Strengths, limitations and future research</b> .....	44
<b>5.5</b>	<b>Ethics</b> .....	45
<b>6.</b>	<b>CONCLUSION</b> .....	45
<b>7.</b>	<b>REFERENCES</b> .....	47
<b>Appendix</b>	.....	60
<b>Papers 1-4</b>		

## **ACKNOWLEDGEMENTS**

The studies reported in this thesis are part of the research project “Emotional and cognitive development of children and adolescents with neuropsychiatric disorders,” which began in 2009. My sincerest gratitude goes to my main supervisor Professor Merete Øie for giving me the opportunity to work on this project. Her extensive knowledge, sharp thinking and enthusiasm for my work has been a tremendous support. I am deeply grateful for her having confidence in me through the entire project. I would also like to thank Professor Jens Egeland, Professor Kenneth Hugdahl and Professor Martin Teicher for their invaluable contributions to papers I, II and III. Thanks to my fellow Phd students Kjell Tore Hovik and Per Norman Andersen for always being positive and helpful when time has been limited and the work load unlimited. As a psychologist student, I learned that discussions during coffee breaks were often as fruitful as many hours of reading. The coffee and bun breaks with Per and Kjell Tore have only reinforced my earlier experience in this area. I would also like to express my deep gratitude to the clinicians, Anne Ringen Amundsen, Audhild Husteli, Berit Dystvold, Eli Gustafsson, Ellen Mælan, Gunhild Winther Skogli, Hanan Almorabet, Heidi Langseth, Ida Enemo, Kjell Tore Hovik, Laila Hagen, Marthe Finstadsveen, Per Normann Andersen, Sigrid Nygaard and Unni Risset for their contribution in collecting baseline data from more than 180 participants in the project. A special thanks also to those helping me and my fellow Phd students with the reassessment of participants in 2011/2012. I am also deeply grateful to all the young participants and their caregivers, who patiently completed our lengthy examinations and assessments. Not only once, but twice, and with the same willingness to share their experiences with us the second time as the first time. Thank you to the Research Department of Innlandet Hospital Trust, and Regional Resource Center for Autism, ADHD, Tourette syndrome and Narcolepsy for financial support, and Lien Diep at Oslo University Hospital for statistical guidance. Thanks to the Department of Psychology at the University of Oslo for letting me follow their Phd program. Thanks to BUP Lillehammer for being an enthusiastic collaborator in this project and for providing me with the necessary practical resources. Last, but not least, I want to thank my parents and sister for the long-lasting support and encouragement to work hard at school, and my wife and two daughters for being patient and supportive at times when the Phd-project has been demanding.

## **SUMMARY**

There is considerable consensus that impairment in executive functions (EF) constitutes a central source of the disability in children and adolescents with ADHD. Recent studies, however, have reported that EF test results are only weakly related to the severity of ADHD symptoms. These findings have led to an increased interest in how EF impairments appear in complex everyday situations where motivational processes interact with EF. Knowledge about this interplay between EF and motivation in children and adolescents with ADHD is scarce. In particular, the degree and type of impairment across ADHD subtypes and genders remains unclear. Therefore, the first aim of this study (paper I) was to compare executive processes with pronounced (hot EF) and less pronounced (cold EF) motivational salience in children and adolescents with ADHD-combined (ADHD-C) and ADHD-inattentive (ADHD-I) subtypes relative to typically developing (TD) children. Our findings supported and expanded on previous studies reporting cold executive dysfunction in children and adolescents with ADHD. There were few differences between ADHD subtypes. Hot EF performance was not impaired in any of the ADHD subtypes relative to TD children.

As most EF tests are highly complex and involve many different sub-processes (both EF and non-EF), low test specificity of the EF tests may hamper the identification of distinct EF profiles in ADHD subtypes. The second aim of the study (paper II) was therefore to investigate cognitive inhibition in a forced attention dichotic listening (DL) task in children with ADHD-C, ADHD-I, and TD children. Results demonstrated impaired cognitive inhibition in both ADHD subtypes, but also significantly greater impairment in the ADHD-I group relative to the ADHD-C group. As far as we know, this study is the first to demonstrate distinct levels of inhibitory impairment between ADHD subtypes in children and adolescents using the DL task.

With regard to gender, previous research has yielded more similarities than differences in cold EF profile for girls and boys with ADHD. However, different symptom expression in boys and girls with ADHD seem to cause referral bias leading to under identification and lack of treatment for females with ADHD. In addition, the presence of co-existing symptoms often clouds the diagnostic picture. The third aim in our study (paper III) was to explore whether there were gender differences in co-existing symptom severity and EF impairment in children and adolescents with ADHD. Self-report and parent rating scales were used to assess co-existing symptoms, whereas EF impairment was assessed with cold EF tests and parent-ratings of EF problems in everyday life. Our results indicate that EF problems in everyday life are more pronounced in boys with ADHD compared to girls with ADHD, despite similar

levels of impairment on cold EF tests. In contrast, co-existing problems are more prominent in girls with ADHD. In particular, self-report scales seem to increase awareness of internalizing problems in females with ADHD. These findings elucidate the combined value of self-report and parent rating scales for the identification of EF problems and comorbid symptoms in boys and girls diagnosed with ADHD.

The fourth aim of the study (paper IV) was to investigate the development of hot and cold EF in boys and girls with ADHD relative to TD counterparts after two years. We found that cold EF remained impaired compared to TD children, despite improved performance across time in the ADHD group. A decline in co-existing problems and ADHD symptoms showed few associations with improved cold EF performance across time. Females with ADHD outperformed TD counterparts on the hot EF task at baseline, but showed deteriorating hot EF performance across time. In contrast, TD females showed improved hot EF performance across time. Our findings indicate that hot and cold EF show divergent developmental trajectories, which are relatively independent from co-existing problems and ADHD symptomatology in children and adolescents with ADHD.

Possible clinical implications of the studies may be that parent ratings of EF seem more suitable than neurocognitive testing to differential diagnosis of ADHD-C and ADHD-I. Further, it may be important to address the issue of techniques to improve cognitive inhibition for children and adolescents with ADHD. Finally, the combination of self-report and parent rating scales are important for the identification of different comorbid symptom expression in boys and girls diagnosed with ADHD.



## LIST OF PAPERS

- I. Skogli, E. W., Egeland, J., Andersen, P. N., Hovik, K. T., & Øie, M. (2013). **Few differences in hot and cold executive functions in children and adolescents with combined and inattentive subtypes of ADHD.** *Child Neuropsychology*, 1-20. doi: 10.1080/09297049.2012.753998
- II. Øie, M., Skogli, E. W., Andersen, P. N., Hovik, K. T., & Hugdahl, K. (2014). **Differences in cognitive control in children and adolescents with combined and inattentive subtypes of ADHD.** *Child Neuropsychology*, 20(1), 38-48.
- III. Skogli, E. W., Teicher, M. H., Andersen, P. N., Hovik, K. T., & Øie, M. (2013). **ADHD in girls and boys -- gender differences in co-existing symptoms and executive function measures.** *BMC psychiatry*, 13(1), 298.
- IV. Skogli, E. W., Andersen, P. N., Hovik, K. T., & Øie, M. (in press). **Development of hot and cold executive function in boys and girls with ADHD: A two-year longitudinal study.** *Journal of Attention Disorders*

## ABBREVIATIONS

ADHD	Attention Deficit Hyperactivity Disorder
ADHD-C	Attention Deficit Hyperactivity Disorder combined subtype
ADHD-I	Attention Deficit Hyperactivity Disorder inattentive subtype
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
BRIEF	Behavior Rating Inventory of Executive Function
CBCL	Child Behavior Checklist
CW 3	Color-Word Interference Test, Condition 3
CW 4	Color-Word Interference Test, Condition 4
DF	Design Fluency Test, Condition 3
D-KEFS	Delis-Kaplan Executive Function System
DL	Dichotic Listening Task
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders - fourth revision
EF	Executive Function
HDT	Hungry Donkey Task
IGT	Iowa Gambling Task
IQ	Intelligence quotient
KSADS-PL	Kiddie-Schedule for Affective Disorders and Schizophrenia Present and Lifetime version
LF	Letter Fluency Test
LN	Letter-Number Sequencing Test
MANOVA	Multivariate Analysis of Variance
RCMAS-2	Revised Children's Manifest Anxiety Scale, second edition
SD	Standard Deviation
SMFQ	Short Mood and Feelings Questionnaire
STAIC	State-Trait Anxiety Inventory for Children
TD	Typical Developing
TMT 4	Trail Making Test, Condition 4
WISC-IV	Wechsler Intelligence Scale for Children – fourth revision
WASI	Wechsler Abbreviated Scale of Intelligence

## 1. INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD) is one of the most common childhood neuropsychiatric disorders. The condition is characterized by problems with inattention, hyperactivity and impulsivity (American Psychiatric Association., 2000). Worldwide prevalence estimates for childhood ADHD range between 3% and 7% (American Psychiatric Association., 2000) with a male-to-female ratio of 3:1 in population based studies (Barkley, 2006; Gaub & Carlson, 1997b). In clinical samples the male-to-female ratio is between 5:1 to 9:1 (American Psychiatric Association., 2000; Gaub & Carlson, 1997b). For more than 50% of the subjects, the difficulties persist into adulthood and have a negative impact on occupational and private life (Barkley & Murphy, 2010; Biederman et al., 2004; Biederman, Monuteaux, et al., 2006). ADHD is believed to be highly heritable with twin studies indicating a heritability of 75% - 77% (Biederman, 2005; Faraone et al., 2005). Despite a strong genetic component, various environmental factors may play a secondary role in the pathogenesis of ADHD. Prenatal exposure to nicotine and alcohol, birth complications, virus infections, meningitis, encephalitis, head injury, epilepsy, toxins, and drug use are associated with increased risk of developing ADHD symptoms (Banerjee, Middleton, & Faraone, 2007; Biederman, 2005; Millichap, 2008). In addition, psychosocial adversity (paternal criminality, maternal mental disorder, foster placement, and low social class) may play a role in the etiology of the disorder (Banerjee et al., 2007; Biederman, 2005; Millichap, 2008).

A general consensus among researchers holds that impairment in executive functions (EF) constitutes a central source of the disability in children and adolescents with ADHD (Seidman, 2006; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Previous research has yielded more similarities than differences in EF profile for ADHD subtypes and for girls and boys with ADHD (Rucklidge, 2010; Seidman et al., 2005; Willcutt et al., 2005). However, with regard to the degree and type of impairment, findings are inconsistent across ADHD subtypes and gender. In particular, the interplay between motivation and EF in naturalistic settings across different ADHD subtypes and genders is unclear. Further, the developmental trajectories of different EF components in boys and girls with ADHD remain unexplored.

In clinical practice, co-existing psychiatric problems are the rule rather than the exception, and the negative impact of having comorbid symptoms on treatment outcome in children with ADHD is well established (Arnold, 1996; Barkley, 2006). Unfortunately, the presence of co-existing symptoms and different ADHD symptom expression in boys and girls with ADHD seem to cloud the diagnostic picture, causing under-identification and lack of treatment for females with ADHD (Quinn, 2008; Rucklidge, 2010). Instead, other diagnoses

are often made long before the diagnosis of ADHD is assessed in female clinical populations (Kopp, Kelly, & Gillberg, 2010).

As part of the research project "Cognitive and emotional development in children and adolescents with neuropsychiatric disorders" at Innlandet Hospital Trust (IHT), the current study aimed to investigate whether EF and emotional regulatory functions differ across diagnostic subgroups and gender in children and adolescents with ADHD. Increased knowledge about the developmental course of EF components and emotional problems across diagnostic subtypes and gender may have relevance for theoretical models of executive dysfunction in ADHD as well as important clinical implications.

Before describing the project, a brief description of the DSM-IV diagnostic criteria for ADHD will be presented (1.1). Further, comorbid problems (1.2) and gender differences (1.3) in children and adolescents with ADHD will be described. Then, executive dysfunction and the interplay between EF and motivational processes will be outlined (1.4). Finally, the developmental course of EF components in TD children and ADHD counterparts will be presented (1.5).

### **1.1 ADHD and diagnostic criteria**

Following the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 2000), ADHD can be divided into three subtypes: the predominantly inattentive subtype (ADHD-I), the predominantly hyperactive/impulsive subtype (ADHD-H) and the combined subtype (ADHD-C). The predominantly inattentive type (ADHD-I) includes individuals with six or more symptoms of inattention and fewer than six symptoms of hyperactivity and/or impulsivity. The predominantly hyperactive-impulsive type (ADHD-H) includes individuals with six or more symptoms of hyperactivity and/or impulsivity and fewer than six symptoms of inattention. The combined type (ADHD-C) is defined by six or more symptoms on both dimensions. These symptoms must: 1) be observed before 7 years of age; 2) manifest themselves across at least two domains; and 3) be chronic. Despite large differences between groups in terms of symptoms, comorbid disorders and neurocognitive impairments, these three subtypes belong to the same diagnostic entity in DSM-IV (Adams, Derefinko, Milich, & Fillmore, 2008; Diamond, 2005). In recent years, more focus has been put on ADHD as a heterogeneous disorder due to large differences within this diagnostic group (Milich, Balentine, & Lynam, 2001; Wahlstedt, Thorell, & Bohlin, 2009).

## **1.2 Comorbidity in ADHD**

As many as 75% of children with ADHD are likely to have at least one other psychiatric disorder (Barkley, 2006; Barkley, Murphy, & Fischer, 2010). In clinical practice, co-existing psychiatric problems are the rule rather than the exception. The presence of co-existing symptoms often clouds the diagnostic picture and may impede correct diagnosis and treatment. Increased understanding of the ability to regulate emotions in children and adolescents with ADHD will be important to understand the mechanisms behind the disease and to provide the best possible treatment. The most common additional difficulties in ADHD are behavioral difficulties and internalizing problems such as anxiety and depression (Biederman, 2005). Behavioral difficulties are more related to ADHD-C than ADHD-I, while most studies have not found consistent differences between ADHD-C and ADHD-I in terms of internalizing symptoms (Gaub & Carlson, 1997a; Willcutt et al., 2012). The negative impact of comorbid symptoms on treatment outcome in children with ADHD is well established (Arnold, 1996; Barkley, 2006).

## **1.3 Gender differences in ADHD**

Females with ADHD are reported to have fewer hyperactive/impulsive symptoms and more inattentive symptoms compared with male counterparts (Biederman et al., 2005; Gaub & Carlson, 1997b; Gershon, 2002). As a result, females with ADHD present more commonly with the inattentive subtype than do boys (Hinshaw, Owens, Sami, & Fargeon, 2006). Different symptom expression, with less disruptive behavior in females with ADHD may contribute to referral bias causing underidentification and lack of treatment for females with ADHD (Biederman et al., 2002; Gaub & Carlson, 1997b; Quinn, 2008; Rucklidge, 2010; Scituito & Eisenberg, 2007). Scituito, Nolfi, & Bluhm (2004) found, for example, that teachers more often refer boys than girls for treatment for ADHD, despite both genders showing equal levels of impairment.

With regard to comorbidity, externalizing problems seem to be more salient in boys with ADHD compared to TD counterparts, whereas internalizing problems are more prevalent in girls with ADHD relative to TD counterparts (Quinn, 2008). Further, adolescent females with ADHD are reported to show more internalizing symptoms than their male counterparts (Rucklidge, 2010). One potential confounding factor with regard to comorbidity profile in different ADHD subtypes may be an unequal distribution of boys and girls in the disorders. As females with ADHD more often present with the inattentive subtype than do boys, this may affect the comorbidity profile in ADHD subtypes (Hinshaw et al., 2006).

#### **1.4 Executive dysfunction – a core feature in children and adolescents with ADHD**

A large number of studies have documented neurocognitive deficits in children and adolescents with ADHD, particularly for executive functions (EF) (Seidman, 2006; Willcutt et al., 2005). Difficulties with EF are associated with poor behavioral, social, educational and occupational outcomes in ADHD populations, and meta analyses have found robust associations between EF impairments and ADHD (Barkley & Murphy, 2010; Biederman et al., 2004; Biederman, Petty, et al., 2006; Seidman, 2006; Willcutt et al., 2005). As behavioral symptoms associated with ADHD have not always proven to be stable over time, EF has been regarded as a promising endophenotype in ADHD (Barkley, 1997; Chhabildas, Pennington, & Willcutt, 2001; Nigg, 2001).

Although there is some debate over exactly which abilities constitute EF, there is growing agreement that EF is a multi-dimensional construct involving higher order cognitive processes such as inhibition, working memory, cognitive flexibility, planning, and verbal fluency (Pennington & Ozonoff, 1996). More recent findings have proposed inhibition, working memory and cognitive flexibility as foundational components in EF (Best & Miller, 2010; Miyake et al., 2000). Due to the multi-dimensionality of the EF construct, no single definition exists. Gioia, Isquith, Guy and Kenworthy (2000, p.1) emphasize the regulation of both cognition, behavior and emotion in their definition of EF: *“The executive functions are a collection of processes that are responsible for guiding, directing, and managing cognitive, emotional, and behavioral functions, particularly during active, novel problem solving.”* The definition proposed by Delis (2012, p.14) on the other hand, emphasizes the multidimensionality of this construct: *“Neither a single ability nor a comprehensive definition fully captures the conceptual scope of executive functions; rather, executive functioning is the sum product of a collection of higher level skills that converge to enable an individual to adapt and thrive in complex psychosocial environments.”* Based on the definitions by Gioia (2000) and Delis (2012), the regulation of cognition, behavior and emotion by multiple cognitive components seems to be essential for EF.

##### **1.4.1 Barkley’s model of a core deficit in response inhibition (1997)**

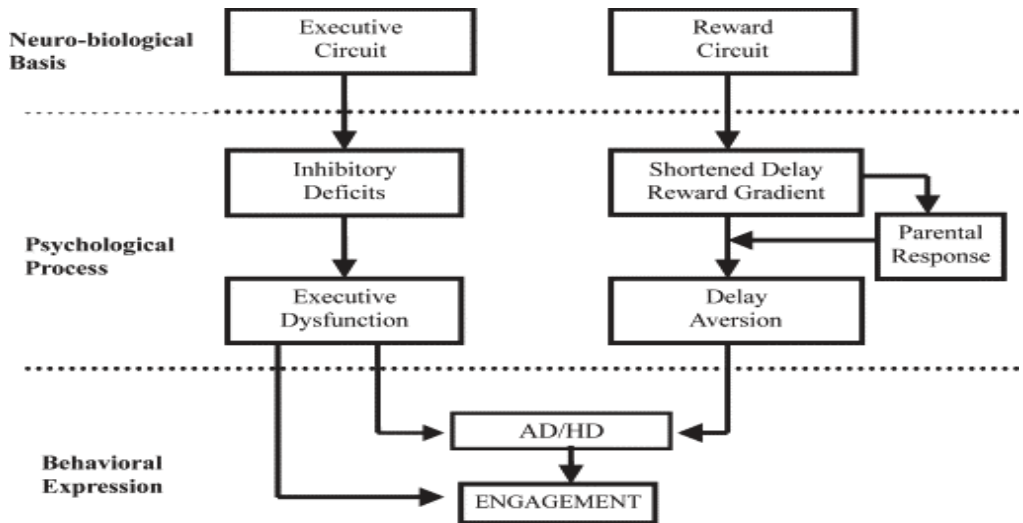
Probably the most influential model of executive dysfunction in ADHD was proposed by Russel Barkley (1997), suggesting that ADHD was characterized by a core deficit in response inhibition. Theoretically, this primary deficit in response inhibition caused secondary EF deficits in working memory, cognitive flexibility, planning and fluency. However, criticism

has been raised against Barkley's core deficit theory as converging evidence does not support the hypothesis that EF deficits are necessary and sufficient causes of ADHD (Nigg, 2005; Willcutt et al., 2005).

#### **1.4.2 Sonuga-Barke's dual pathway model of ADHD (2003)**

The growing consensus that the EF "single deficit" model is insufficient to explain ADHD symptoms has led to a shift from the core deficit theory to multiple deficits theories. According to Sonuga-Barke's (2003, 2005) dual pathway model of ADHD, two separate neurodevelopmental pathways can lead to ADHD. The first "executive dysfunction" pathway associated with abnormalities in dorsolateral circuits causes cold EF deficits (Barkley, 1997; Sonuga-Barke, 2003, 2005). The second "motivational dysfunction" pathway associated with abnormalities in orbitofrontal circuits has been linked to suboptimal reward processes in children with ADHD (Sagvolden, Aase, Zeiner, & Berger, 1998). Aberrant levels of sensitivity to rewards and penalties in children and adolescents with ADHD may cause greater preference for immediate over delayed rewards compared to typical developing (TD) children. Likewise, children and adolescents with ADHD show less sensitivity to negative feedback, and seem to over-respond to immediate reward (see Luman, Oosterlaan, & Sergeant, 2005 for a review). As proposed by the dual pathway model, these suboptimal reward processes may independently cause ADHD symptoms, eventually in conjunction with impaired EF (Sonuga-Barke, 2003, 2005).

#### **Figure 1: The Dual Pathway Model**



Sonuga-Barke (2003)

### 1.4.3 Cold EF in children and adolescents with ADHD

Traditionally, most neurocognitive tests measuring EF reflect cognitive processes with little emotional salience, typically referred to as cold EF (Chan, Shum, Touloupoulou, & Chen, 2008). These cold EF tests demand relatively simple responses in a quiet and structured test environment. They may assess maximum performance, but not always a typical function (Toplak, West, & Stanovich, 2013). In Barkley's (1997) model, cold EF impairments are explained by a primary deficit in cognitive inhibition and are related to ADHD-C but not to ADHD-I. With regard to the ADHD-I group, Diamond (2005) argues in a comprehensive review that working memory impairments constitute a core deficit. However, in a meta-analytic review of 83 studies with sufficient group sizes and numerous EF tasks, Willcutt, Doyle, Nigg, Faraone, and Pennington (2005) did not find support for reliable EF differences between the ADHD-C and ADHD-I subtypes. Overall, the research on cold EF deficits is inconsistent, making ADHD subtype differentiation unreliable. With regard to gender, previous research has yielded more similarities than differences in cold EF profiles for girls and boys with ADHD (Rucklidge, 2010; Seidman et al., 2005).

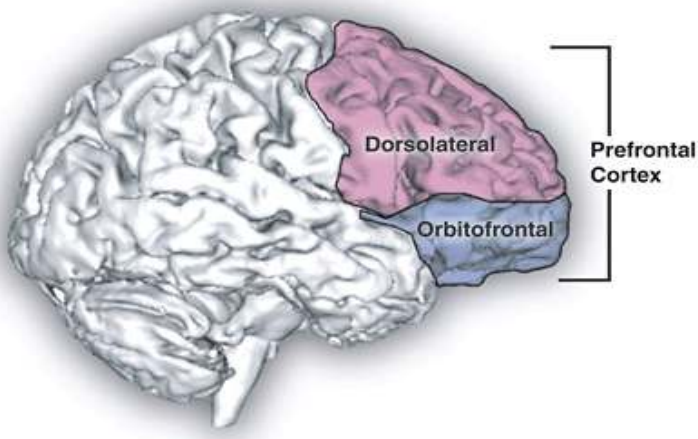
### 1.4.4 Hot EF in children and adolescents with ADHD

Even though cold EF has been regarded as a promising endophenotype in ADHD, cold EF tests have shown to be weakly related to the severity of ADHD symptoms (Chhabildas et al.,



2001; Jonsdottir, Bouma, Sergeant, & Scherder, 2006; Nigg, 2001). Consequently, there has been increasing interest in how EF impairments often appear in complex everyday situations where emotions and motivation interact with cold EF processes (Sonuga-Barke, 2003, 2005). Therefore, deficits in cold EF and emotional-reward dysregulation may constitute independent routes to ADHD symptoms in general and represent potential loci for ADHD subtype characteristics in particular (Castellanos, Sonuga-Barke, Milham, & Tannock, 2006; Sonuga-Barke, 2003, 2005). Executive processes in situations with stronger affective salience are often referred to as hot EF. Hot EF processes have shown to activate areas of the brain that regulate emotions and the brain's reward systems (e.g., orbito-frontal cortex, ventral striatum, and the limbic system), whereas traditional cold EF tests activate the dorsolateral parts of the prefrontal cortex (Castellanos et al., 2006). As proposed by multiple pathway models in ADHD, hot EF processes may constitute another promising endophenotype explaining ADHD symptoms (Sonuga-Barke, 2003, 2005).

Figure 2: Prefrontal cortex



Source: mybrainnotes.com

#### 1.4.5 EF in everyday situations in children and adolescents with ADHD

Whereas traditional neurocognitive EF tests seem to capture “best estimates” in an ideal setting (McCue & Pramuka, 1998), the Behavior Rating Inventory of Executive Function (BRIEF) was designed to assess EF performance in ecologically valid situations (Gioia, Isquith, Guy, &

Kenworthy, 2000). In clinical practice, the BRIEF has proven to be a valuable additional assessment tool for the identification of ADHD in school-aged children, demonstrating better sensitivity in ADHD assessment than neurocognitive EF tests (McCandless & O'Laughlin, 2007; Toplak, Bucciarelli, Jain, & Tannock, 2009; Toplak et al., 2013). However, rather modest correlations between BRIEF and neurocognitive EF tests indicate that the BRIEF and conventional neurocognitive EF tests are complimentary rather than similar instruments (V. A. Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Bodnar, Prahme, Cutting, Denckla, & Mahone, 2007; Toplak et al., 2009).

### **1.5 Typical development of EF**

One of the challenges in understanding hot and cold EF in children and adolescents is that these skills develop and change rapidly throughout childhood (P. Anderson, 2002; Best & Miller, 2010). Further, it appears that different EF components demonstrate divergent developmental trajectories, with persistent maturation of some skills, and developmental delay reported for other EF components (e.g. decision-making) (P. Anderson, 2002; Smith, Xiao, & Bechara, 2012). In TD subjects, cold EF are considered to be fully matured in early adulthood, with EF components cognitive inhibition and working memory developing ahead of cognitive flexibility (Best & Miller, 2010). This protracted maturation of EF is illustrated by structural imaging findings (e.g., using MRI) showing that the prefrontal cortex continues to develop (continued white matter development due to myelination) through adolescence, unlike many other brain regions (Faraone et al., 2005; O'Hare, Sowell, Nelson, & Luciana, 2008). Adding further complexity to an understanding of EF development in children and adolescents with ADHD, growing evidence indicates that the brains of TD boys and girls develop and mature at different rates (Lenroot et al., 2007; Thompson et al., 2005). Magnetic resonance imaging (MRI) studies have reported that frontal lobes pivotal for EF mature 1-3 years earlier in TD girls compared to TD boys (Giedd et al., 2009; O'Hare et al., 2008).

#### **1.5.1 Development of cold EF in children and adolescents with ADHD**

Despite improved cold EF performance through childhood in children with ADHD, cold EF seem to remain significantly impaired into adolescence and adulthood (Seidman, 2006). Further, inattention symptoms are reported to be more closely related to EF deficits than hyperactivity/impulsivity symptoms in children and adolescents with ADHD (Sonuga-Barke, 2005; Willcutt et al., 2012). With regard to gender, previous studies have reported comparable developmental trajectories in cold EF for boys and girls with ADHD (Seidman, 2006).

### **1.5.2 Development of hot EF in children and adolescents with ADHD**

One of the most widely used measures of hot EF is the Iowa Gambling Task (IGT) (Bechara, Damasio, Damasio, & Anderson, 1994). On a decision-making task like the IGT and its child friendly analogue the Hungry Donkey Task (HDT) (Crone & van der Molen, 2004), the interaction between hot and cold EF is essential to make long-term, advantageous choices. Prencipe and colleagues (2011) proposed that age-related improvements occur later and are more gradual for hot executive processes than for cold EF in TD children. This is supported by findings that advantageous decision-making strategies have only been reported for older groups of children (>12 years) in both clinical and TD populations (Crone & van der Molen, 2004; da Mata et al., 2011; Overman, 2004; van Duijvenvoorde, Jansen, Bredman, & Huizenga, 2012). Thus, as the ability to make long-term advantageous choices increases with age in TD adolescents, adolescents with ADHD may remain impaired. This may explain why impaired decision-making has been found to be more pronounced in adolescents with ADHD than in children with ADHD (Geurts, van der Oord, & Crone, 2006; Hobson, Scott, & Rubia, 2011; Lambek et al., 2010; Toplak, Jain, & Tannock, 2005). Further, previous findings indicate that there may be sex-related differences in decision-making performance, with males outperforming females in TD children and in adults (Crone, Bunge, Latenstein, & van der Molen, 2005; Garon & Moore, 2007; Overman, 2004; van den Bos, Homberg, & de Visser, 2013). In a review by van den Bos and colleagues (2013), the authors suggested that gender differences on decision-making tasks were caused by different strategies in males and females. Whereas males in general focus on the overall outcome, females to a greater extent focus on both win and loss frequencies as well as overall outcome. As a consequence, females tend to be more risk averse and need more trials to reach the same level of performance as males (van den Bos et al., 2013). Few studies have investigated the impact of gender on decision-making in ADHD populations, but Miller and colleagues (2013) found impaired decision-making performance in young adult females diagnosed with ADHD relative to TD females. Despite the lack of between sex comparisons in this study, results indicate that young adult females with ADHD exhibit impaired decision-making, similar to their male counterparts.

### **1.6 Unanswered questions**

There is considerable consensus that EF impairments constitute a central source of disability in children and adolescents with ADHD (Seidman, 2006; Willcutt et al., 2005). However,

with regard to the degree and type of impairment, the findings are inconsistent across ADHD subtypes and gender. In particular, the impact of hot EF and emotional regulation in naturalistic settings across ADHD-I/ADHD-C and gender remains unclear. Further, the longitudinal course of hot EF in boys and girls with ADHD is an understudied topic.

## **2. AIMS**

The main objective of the present study was to investigate EF and emotion regulation across subtypes and gender in children and adolescents with ADHD. This was explored cross-sectional and longitudinal.

In paper I and II, the aims were to investigate whether hot and cold EF measures could differentiate between children and adolescents with ADHD-I and ADHD-C subtypes. Neurocognitive EF profiles were compared between groups of children with ADHD-I and ADHD-C.

In paper III, the aim was to delineate specific ratings and measures of EF abilities that were most important in distinguishing ADHD subjects from TD subjects of the same gender. Differences in co-existing symptom severity by gender were also explored.

In paper IV, the aim was to investigate the development of hot and cold EF across time in boys and girls with ADHD relative to TD children. Potential gender effects were also examined.

## **3. METHODS**

### **3.1 Design**

Varying designs were used in the study. Papers I, II and III were based on a naturalistic, cross-sectional design. Paper IV was based on a naturalistic longitudinal design. The study is part of the ongoing research project “Cognitive and emotional development in children and adolescents with neuropsychiatric disorders”, aiming to investigate cognitive and emotional regulatory functions in children and adolescents with various neuropsychiatric disorders. The subjects were recruited as consecutive referrals for assessment of ADHD from all seven Child and Adolescent outpatient Mental Health Centres in Hedmark and Oppland. The TD subjects were recruited from local schools in the same catchment area. Clinical and neurocognitive data were collected from all participants. All ADHD participants received medication and/or

standard psychological treatment in the period between inclusion and reassessment after two years.

### **3.2 Procedure**

All participants underwent a comprehensive assessment according to common clinical practice. Neurocognitive assessment was carried out at the same time as clinical evaluation, and included the administration of a three hour long neurocognitive battery. All neurocognitive testing included two breaks of 15 minutes. The interviewers were experienced clinicians, and all were trained in neurocognitive testing and diagnostic assessment prior to inclusion of participants. The project manager, who is a clinical specialist in neuropsychology, independently reviewed all diagnostic assessments set by the clinicians. Meetings between the clinicians and the project manager were held regularly to arrive at a consensus when the diagnostic assessment was inconclusive. The TD children responding positively to the letter inviting them to participate received a phone call for time for assessment. All TD participants underwent the same assessment procedures as the clinical participants.

Follow-up (T2) assessment was conducted approximately 24 months following baseline assessment. The same procedure was followed in both assessments. Seventy-seven subjects with ADHD from the baseline sample of 80 were available for reassessment at T2 (96%). Three boys with ADHD refused to participate at T2. In four cases where the participants were unable to meet at the clinic for reassessment, telephone interviews were performed with the parents/caregivers. Of the 77 subjects with ADHD available for reassessment, 2 boys and 2 girls were symptom free and received no diagnosis at T2. Due to fatigue or refusal to perform some of the tests, a complete neurocognitive re-assessment was not possible in all cases. In one case, computer failure precluded full data collection (i.e., the Hungry Donkey Task).

### **3.3 Participants**

At baseline, 80 subjects with ADHD and 50 TD children ages 8 to 17 years participated in the study. The mean age was 11.5 ( $SD = 2.0$ ) in the ADHD group and 11.6 ( $SD = 2.0$ ) in the TD group. Paper I and III included the same sample. In paper II, the TD group was replaced with age and sex-matched TD subjects from the Bergen study. In paper 4, the age span was restricted to 9-16 years instead of 8-17 in order to get greater age homogeneity. One boy and

one girl with ADHD, and three boys in the TD group were excluded due to age restriction criteria in paper IV.

In paper I, there were two diagnostic subgroups: 44 with ADHD-I, and 36 with ADHD-C, in addition to the TD subjects.

In paper II, there were two diagnostic subgroups: 42 with ADHD-I, and 29 with ADHD-C, in addition to 40 age and sex matched TD subjects.

In paper III, four groups were defined. Group 1 consisted of 43 males with ADHD. Group 2 consisted of 37 females with ADHD. Group 3 consisted of 32 TD males. Group 4 consisted of 18 TD females.

In paper IV, four groups were defined. Group 1 consisted of 39 males with ADHD. Group 2 consisted of 36 females with ADHD. Group 3 consisted of 29 TD males. Group 4 consisted of 18 TD females.

All participants with a history of central nervous system pathology, prematurity (< 36 weeks), a history of stimulant treatment or with an estimated IQ less than 70 were excluded. Participants in the TD groups were screened for mental disorders with the Kiddie-Schedule for Affective Disorders and Schizophrenia in separate interviews for children/adolescents and parents. All TD participants with a history of psychopathology, head trauma (with loss of consciousness), or dyslexia were excluded from the study.

### **3.4 Measures**

#### **3.4.1 Clinical assessment**

Diagnosis was determined based on semi-structured clinical interviews and standardised rating scales. The interview using the Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADS) (Kaufman et al., 1997) was conducted separately for children/adolescents and parents to assess psychopathology. The diagnostic evaluation with K-SADS was supplemented with information from the ADHD Rating Scale IV (ARS-IV) (DuPaul, Power, Anastoupoulos, & Reid, 1998), and the Child Behavior Checklist/6-18 (Achenbach & Rescorla, 2001). Additional information about school functioning, which is mandatory on referral, was incorporated into the diagnostic evaluation. Diagnoses were considered positive if, based on a comprehensive evaluation of K-SADS, teacher information and standardised rating scales, DSM-IV (American Psychiatric Association., 2000) criteria were met.

### 3.4.2 Neurocognitive assessment

A neurocognitive test battery assessing functions sensitive to EF impairment in ADHD was used (Seidman, 2006; Willcutt et al., 2005). The EF measures included were aimed at assessing working memory, cognitive inhibition, cognitive flexibility, planning, verbal fluency, and decision-making. However, a cornerstone in the understanding of EF is that different EF functions are both correlated (unity) and specific (diversity) (Miyake & Friedman, 2012). Although the tests load on a range of different EF components, we assigned the tests to the following EF domains:

*Working memory:* The Letter-Number Sequencing Test (LN) (Wechsler, 2004) was used as a measure of working memory in papers I, III and IV. The test consists of ten items. Each item contains three trials with the same number of digits and letters. The test administrator reads aloud each trial and asks the child to recall the numbers in ascending order and the letters in alphabetical order. In the present study, total correct recalled trials were examined. Lower scaled scores indicated difficulties with the task. Mean reliability for the LN test across all ages is 0.75 (Wechsler, 2004).

*Cognitive inhibition:* The Color – Word Interference Test, Condition 3 (CW 3) (D-KEFS; Delis, Kaplan, & Kramer, 2001; Stroop, 1935) was used as a measure of cognitive inhibition in papers I, II, and IV. The examinee needs to inhibit an overlearned verbal response when naming the dissonant ink colors in which the words are printed. For the present study, completion time in seconds was examined. Lower scaled scores indicated difficulties with the task. Mean reliability for the CW 3 test ages 8-19 is 0.90 (Delis et al., 2001).

The Dichotic Listening Task (DL) (Hugdahl & Andersson, 1986) is an auditory test of brain laterality and cognitive inhibition (controlled attention). The test was used for the study in paper II. Stimulus materials consisted of six stop-consonants (b, d, g, p, t, and k) that were paired with the vowel a to form six basic consonant-vowel (CV) syllables: ba, da, ga, pa, ta, ka (Hugdahl & Andersson, 1986). The subject hears pair-wise presentations of the consonant-vowel speech sounds and is required to focus attention on either the right or left ear stimulus (Hugdahl & Andersson, 1986). The DL task was administered under three conditions: non-forced (NF) condition, forced-right (FR) condition, and forced-left (FL) condition. Responses were scored as number of correctly reported syllables presented either to the left or right ear during each of the three attentional instructions. The DL test with CV stimuli has yielded test-retest reliabilities of between 0.70 (Bakker, van der Vlugt, & Claushuis, 1978) and 0.90 (Harper & Kraft, 1986).

*Cognitive flexibility:* The Color – Word Interference Test, Condition 4 (CW 4) (D-KEFS; Delis et al., 2001) was used as a measure of cognitive flexibility in paper III. The examinee is asked to switch back and forth between naming the dissonant ink colors and reading the words. For the present study, completion time in seconds was examined. Lower scaled scores indicated difficulties with the task. Mean reliability for the CW 4 test ages 8-19 is 0.80 (Delis et al., 2001).

The Trail Making Test, Condition 4 (TMT 4) (D-KEFS; Delis et al., 2001) was used as a measure of cognitive flexibility in papers I, III, and IV. The examinee is asked to draw a line interchangeably between numbers and letters in the right order. For the present study, time to complete task was examined. Lower scaled scores indicated difficulties with the task. Mean reliability for the TMT 4 test ages 8-19 is 0.20 (Delis et al., 2001).

The Design Fluency Test, Condition 3 (DF) (D-KEFS; Delis et al., 2001) was used as a measure of cognitive flexibility in paper III. The examinee is asked to draw as many different designs as possible using four straight lines connecting five filled and empty dots interchangeably. The examinee is given 60 seconds for the task. For the present study, total correct responses were examined. Lower scaled scores indicated difficulties with the task. Mean reliability for the DF test ages 8-19 is 0.13 (Delis et al., 2001).

*Planning:* The Tower Test (D-KEFS; Delis et al., 2001) was used as a measure of planning abilities in papers I and III. In this task, the examinee is asked to construct several target towers by moving five disks, varying in size, across three pegs in the fewest number of moves possible. While doing this, the examinee is allowed to move only one disk at a time, and not to place a larger disk over a smaller disk. In the present study, total achievement score within the time limit as specified in the D-KEFS manual was examined. Lower raw scores indicated difficulties with the task. Mean reliability for the Tower test ages 8-19 is 0.51 (Delis et al., 2001).

*Verbal Fluency:* The Letter Fluency Test (LF) (D-KEFS; Delis et al., 2001) was used as a measure of verbal fluency in papers I and III. This task includes three 60-second trials, where participants were asked to generate words fluently in an effortful, phonemic format with the letters F, A, and S. For the present study, total correct responses were examined. Lower scaled scores indicated difficulties with the task. Mean reliability for the LF test ages 8-19 is 0.67 (Delis et al., 2001).

*Decision-making:* The computer-based Hungry Donkey Task (HDT) (Crone & van der Molen, 2004) was used as a measure of hot EF in papers I and IV. Participants were asked to help a hungry donkey collect as many apples as possible by choosing one of four doors (A, B,



C, D). The amount of wins and losses varied between choices, and the overall gains/losses were displayed with a red/green bar at the bottom of the screen. The task ended after completion of 150 trials. Net difference scores were calculated by subtracting the amount of disadvantageous choices (A+B) from the number of advantageous choices (C+D) (e.g. Bechara et al., 1994). Higher net scores indicated better performance on the task. Pearson correlation between odd and even blocks revealed adequate internal consistency on the task in our sample at T1 (.694) and T2 (.628).

*General cognitive functioning (IQ)*: The Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999) was administered to estimate IQ in all participants.

### **3.4.3 Parent-rated EF**

EF impairment in everyday situations was assessed with the *Behavior Rating Inventory of Executive Function (BRIEF)* (Gioia et al., 2000). The version designed for children and adolescents ages 5 to 18 includes a parent form and a teacher form (Gioia et al., 2000). In the current study, the Norwegian parent rating version was used. The BRIEF is composed of eight clinical scales (Inhibition, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials and Monitor). The BRIEF scales Inhibit, Shift, and Emotional Control constitute the broader Behavioral Regulation Index. The BRIEF scales Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor constitute the broader Metacognition Index. The Global Executive Composite is a summary measure derived from all clinical scales (Gioia et al., 2000). Fallmyr & Egeland (2011) reported high internal consistency (Chronbachs  $\alpha = .76-.92$ ) for the Norwegian parent rating version of the BRIEF. These values are at the same level as Chronbachs  $\alpha$  reported in the BRIEF manual (.80-.98) (Gioia et al., 2000). Elevated BRIEF T-scores indicate a higher degree of impairment. A short description of the eight clinical scales are as follows:

*Inhibit*: The ability to control his or her impulses, to stop own behavior at the proper time, and to be able to continue in the right direction and solve the task (Gioia et al., 2000).

*Shift*: The ability to move freely from one situation, activity, or aspect of a problem to another in a given situation (Gioia et al., 2000).

*Emotional Control*: The ability to modulate emotional responses appropriately (Gioia et al., 2000).

*Initiate*: The ability to begin a task or activity, and independently generate ideas that are helpful in solving a specific task or problem (Gioia et al., 2000).

*Working memory:* The ability to hold information in mind for the purpose of completing a task (Gioia et al., 2000).

*Plan/Organize:* The ability to manage current tasks and to anticipate future goals (Gioia et al., 2000).

*Organization of Materials:* The ability to organize the world (e.g. workspace, play areas) and keep order of the belongings (Gioia et al., 2000).

*Monitor:* The ability to check work and the effect of own behavior on others (Gioia et al., 2000).

### **3.4.4 Symptom assessment**

#### 3.4.4.1 Parent-rating scales

*The Child Behavior Checklist/6-18 (CBCL)* (Achenbach & Rescorla, 2001) is a widely used scale containing 7 competence items and 113 specific problem items, each of which is rated on a 0–2 metric. The 120 items assess adaptive behavior as well as eight narrow band factors (Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior and Aggressive Behavior) and two broadband factors (Externalizing and Internalizing symptoms) of co-existing symptoms. Elevated T-scores indicate a higher degree of co-existing internalizing and externalizing symptoms. Cross-cultural studies have demonstrated good discriminant validity with mean factor loadings across societies at .62 (Ivanova et al., 2007). Acceptable reliability and validity of the Norwegian version of the CBCL are reported by Nøvik (Nøvik, 1999, 2000).

#### 3.4.4.2 Self-report scales

*The Revised Children's Manifest Anxiety Scale, second edition (RCMAS-2)* (Reynolds & Richmond, 1985) is a 49-item self-report instrument designed to measure anxiety symptoms in children 6 to 19 years of age. Children respond either “Yes” or “No” to all 49-items. The instrument reveals three anxiety factors: Physiological Anxiety, Worry and Social Anxiety. The three anxiety factors are summed yielding a Total Anxiety score. Elevated raw-scores indicate a higher degree of anxiety symptoms. The RCMAS Total Anxiety Scale has been found to have satisfactory psychometric properties with high test–retest reliability (Pela & Reynolds, 1982; Reynolds, 1981) and consistent construct validity (Reynolds, 1980, 1982; Reynolds & Paget, 1981; Scholwinski & Reynolds, 1985). Satisfactory psychometric properties have been replicated among other cultures as well (Boehnke, Sillbereisen,

Reynolds, & Richmond, 1986; Ferrando, 1994; Pela & Reynolds, 1982; Turgeon & Chartrand, 2003).

*The State-Trait Anxiety Inventory for Children* (STAIC) includes two 20-item self-report scales that measure both enduring tendencies (Trait) and situational variations (State) in levels of perceived anxiety. Children respond on a three-point scale indicating varying degree of worry, feelings of tension, and/or nervousness. Elevated raw-scores indicate a higher degree of situational and temporal anxiety. In a quantitative review by Seligman and colleagues (2004), the authors argue that the STAIC possess satisfactory psychometric properties.

*The Short Mood and Feelings Questionnaire* (SMFQ) is a 13-item self-report instrument designed to measure depressive symptoms in children and adolescents. The SMFQ is derived from the original 30-item Mood and Feelings Questionnaire (MFQ) (Costello & Angold, 1988) where children respond on a three-point scale (“not true”, “sometimes true” and “true”). A net score was generated based on the 13 items with elevated raw-scores indicating a higher degree of depression symptoms. The SMFQ have demonstrated high internal consistency (Cronbach’s  $\alpha = .90$ ) (Costello, Benjamin, Angold, & Silver, 1991), and test-retest stability in children for a two-week period yielded an intra class correlation of .66 (Costello & Angold, 1988). Angold and colleagues (1995) found SMFQ to correlate strongly with Children’s Depression Inventory (CDI) (Kovacs, 1983) and Diagnostic Interview Schedule for Children (DISC-C) depression scores (Costello & Angold, 1988) ( $r = .67$  and  $.51$ , respectively). Acceptable validity of the Norwegian version of the SMFQ are reported by Lundervold and colleagues (2013).

### **3.5 Statistical analyses**

Data analyses were conducted using the statistical package SPSS for Windows, version 15.0 and 19.0 (SPSS, Inc., Chicago, IL). Demographic characteristics were investigated using the Chi-square test for independence (nominal variables) t-tests (continuous variables) and analysis of variance (ANOVA). All analyses were two-tailed with a significance level of 0.05. Correlations between variables were explored with Pearson’s  $r$ .

Due to the large number of analyses performed in the first paper, multivariate analysis of variance (MANOVA) was conducted in order to control for type I errors. Significant group effects ( $p < .05$ ), were followed up with analyses of variance (ANOVA) and analyses of covariance (ANCOVA). Scheffe post-hoc tests for group comparisons were conducted when relevant. Effect sizes were calculated by  $Eta^2$ .

The dichotic listening data in paper II were subjected to a factorial ANOVA based on the design 3 Groups (ADHD-C, ADHD-I, TD ) x 2 Ear x 3 Attention instruction (NF, FR, FL). Main-effects of group were follow-up with Fischer's LSD test.

In paper III Random forest classification analyses were performed in addition to conventional ANOVA's. Random forest classification which is an algorithmic modelling / data mining technique was used to explore gender differences in co-existing symptoms and EF ratings and measures. Classical statistical techniques are designed to test and reject the hypothesis that observed differences between groups have occurred by chance. Algorithmic modelling techniques have been developed to address a somewhat different question. Briefly, these techniques can identify from a sample of potential predictor variables the most important subset for categorizing subjects or predicting outcomes (Breiman, 2001). Hence, we used this approach to delineate within each gender the subset of symptom ratings, EF measures and EF ratings, that appear to be most important in discriminating children with ADHD from TD children. Specifically, we used random forest classification and cross-validation (R packages *randomForest 4.5-34* and *caret 5.02-011*) (Breiman & Cutler, 2003) to identify and rank order different symptom ratings and EF measures for their degree of importance in differentiating ADHD from TD children within each gender. Random forest classification is a form of "ensemble learning" in which a large number of un-pruned decision trees are generated and their results aggregated. The random part comes in as each tree is constructed using a different bootstrap sample of the data, and each node is split using the best among a subset of predictors randomly chosen at that node. Each decision tree was generated using results from 75% of the participants and then tested on the remaining 25% (validation set). This process was performed 5000 times on different random splits of the data to provide a cross-validated estimate of the predictive discriminant ability of the measures (accuracy, kappa) that would likely generalize to new cases (Kuhn, 2008). The importance of each variable in the cluster was assessed by calculating the decrease in predictive accuracy following the sequential permutation (effective randomization and elimination) of each variable in the cluster on the validation set. The most important variables were the ones whose effective elimination from the forest produced the greatest degradation in accuracy.

In paper IV baseline and follow-up data were analysed longitudinally. Mixed between-within subjects ANOVAs (mixed ANOVA) were conducted for each dependent variable to estimate the effect of group and gender on EF performance across time. Because of multiple comparisons, Bonferroni-Holm corrections were used to control for chance findings by reducing the global alpha level ( $\alpha = .05$ ) proportionately to the number of comparisons being

performed (Holm, 1979). To control for the possible confounding effect of IQ, mixed ANOVAs for all dependent measures with IQ as a covariate were conducted. All significant hot and cold EF test results were re-analysed without the one participant taking stimulant medication prior to testing at T2. Correlation analyses (Pearson) were used to investigate associations between improvement in EF performance between T1 and T2, and the amount of change in co-existing problems and ADHD symptomatology.

The statistical analyses used in the four studies, are described in more detail in the four papers.

#### **4. SUMMARY OF PAPERS**

##### **Paper I: Few differences in hot and cold executive functions in children and adolescents with combined and inattentive subtypes of ADHD.**

Background: The aim of the study was to compare executive processes with pronounced (hot) and less pronounced (cold) emotional salience in medication naïve children and adolescents with ADHD-combined (ADHD-C) and ADHD-inattentive (ADHD-I) subtypes.

Method: Thirty-six subjects with ADHD-C, 44 with ADHD-I, and 50 TD subjects between 8 and 17 years were assessed with laboratory tests and inventory based scales. The test battery included measures assessing EF (cognitive inhibition, working memory, planning, cognitive flexibility, verbal fluency, hot decision-making) and the Behavior Rating Inventory of Executive Function (BRIEF).

Results: The ADHD-C group displayed significantly more impairment compared to the ADHD-I group on the cold BRIEF Inhibition and Monitor scales. There were no significant differences between ADHD subtypes on cold and hot laboratory tests. The hot decision-making task did not correlate with the other cold or hot EF measures.

Conclusion: Overall, few EF measures were shown to differentiate between ADHD subtypes; nor were there any relationships between the hot decision-making task and the other EF measures, which seems to indicate separate developmental trajectories.

##### **Paper II: Differences in cognitive control in children and adolescents with combined and inattentive subtypes of ADHD.**

Background: The aim of the present study was to investigate the ability of children with Attention-Deficit/Hyperactivity Disorder Combined subtype (ADHD-C) and Predominantly Inattentive subtype (ADHD-I) to direct their attention and exert cognitive control in a forced instruction dichotic listening (DL) task.

Method: Twenty-nine, medication naïve subjects with ADHD-C, 42 with ADHD-I and 40 matched TD subjects between 10 and 16 years were assessed. In the DL task, two different auditory stimuli (syllables) are presented simultaneously, one in each ear, without the subject being aware of the dichotic nature of the stimulus presentation. The subjects are asked to report the syllable they hear on each trial with no instruction of focus of attention, or to explicitly focus attention and report either the right- or left ear syllable. The DL procedure is presumed to tap distinct cognitive processes: perception (non-forced condition/NF), orienting of attention (forced-right condition/FR), and cognitive control (forced-left condition/FL).

Results: As expected, all three groups had normally perception and orienting of attention. Children and adolescents with ADHD-I showed a right ear advantage also during the FL condition compared to subjects with ADHD-C and TD participants.

Conclusion: The results may indicate a deficit in cognitive control in the ADHD-I group, presumably mediated by a deficit in prefrontal neuronal circuitry. Our results may have implications for further conceptualization, diagnostics, and treatment of ADHD subtypes.

### **Paper III: ADHD in girls and boys – gender differences in co-existing symptoms and executive function measures**

Background: ADHD is diagnosed and treated more often in males than in females. Research on gender differences suggests that girls may be consistently underidentified and underdiagnosed because of differences in the expression of the disorder among boys and girls. One aim of the present study was to assess in a clinical sample of medication naïve boys and girls with ADHD, whether there were significant gender x diagnosis interactions in co-existing symptom severity and EF impairment. The second aim was to delineate specific symptom ratings and measures of EF that were most important in distinguishing ADHD from TD children of the same gender.

Methods: Thirty-seven females with ADHD, 43 males with ADHD, 18 TD females and 32 TD males between 8 and 17 years were included. Co-existing symptoms were assessed with self-report scales and parent ratings. EF was assessed with parent ratings of executive skills in everyday situations (BRIEF), and neurocognitive tests. The three measurement domains (co-existing symptoms, BRIEF, neurocognitive EF tests) were investigated using analysis of variance (ANOVA) and random forest classification.

Results: ANOVAs revealed only one significant diagnosis x gender interaction, with higher rates of self-reported anxiety symptoms in females with ADHD. Random forest classification indicated that co-existing symptom ratings was substantially better in distinguishing subjects

with ADHD from TD in females (93% accuracy) than in males (86% accuracy). The most important distinguishing variable was self-reported anxiety in females, and parent ratings of rule breaking in males. Parent ratings of EF skills were better in distinguishing subjects with ADHD from TD in males (96% accuracy) than in females (92% accuracy). Neurocognitive EF tests had only a modest ability to categorize subjects as ADHD or TD in males (73% accuracy) and females (79% accuracy).

Conclusions: Our findings emphasize the combination of self-report and parent rating scales for the identification of different comorbid symptom expression in boys and girls already diagnosed with ADHD. Self-report scales may increase awareness of internalizing problems particularly salient in females with ADHD.

#### **Paper IV: Development of hot and cold executive functions in boys and girls with ADHD: A two year longitudinal study.**

Objective: Investigate the development of EF with pronounced emotional salience (hot EF) and less pronounced emotional salience (cold EF) in boys and girls with ADHD relative to TD children.

Method: Seventy-five children with ADHD, and 47 TD children were assessed with hot and cold EF tests at baseline and after two years.

Results: Despite considerable maturation, the ADHD group remained impaired on all cold EF tests relative to TD children after two years. There was no effect of gender on cold EF test results. Females with ADHD outperformed TD counterparts on hot EF at baseline. Females with ADHD showed deteriorating hot EF performance, while TD counterparts showed improved hot EF performance across time.

Conclusion: Enduring cold EF impairments after two years may reflect stable phenotypic traits in children with ADHD. Results indicate divergent developmental trajectories of hot EF in girls with ADHD relative to TD counterparts.

## **5. DISCUSSION**

First, the main findings in the four papers will be discussed in light of previous and recent research (5.1). Second, relevant methodological concerns will be presented (5.2), along with theoretical and clinical implications of our findings (5.3). Strengths and limitations of the study and suggestions for future research will be outlined (5.4), before a few ethical considerations are presented (5.5).

## **5.1 Main findings**

### **5.1.1 Hot and cold EF in children and adolescents with ADHD subtypes**

In the first study we found that few hot and cold EF measures were shown to differentiate between ADHD subtypes. As expected, we found that children and adolescents with either ADHD subtype showed significantly more impaired functioning on working memory and verbal fluency compared to TD children. Only the ADHD-I group displayed significantly more impaired scores on measures of cognitive inhibition and cognitive flexibility relative to TD children. Relative to TD children, the ADHD-I group show a somewhat more widespread impairment in cold EF than the ADHD-C group. It is nevertheless notable that the magnitude of cold EF impairments working memory and verbal fluency in the two ADHD subgroups are equal, as illustrated by comparable effect sizes. Our results are consistent with a meta-analytic review of 83 studies by Willcutt and colleagues (2005), reporting significant impairment across different EF domains (i.e., cognitive inhibition, working memory, cognitive flexibility) in children and adolescents with ADHD, but no reliable EF differences between ADHD-C and ADHD-I subtypes.

As expected, we found no significant differences between groups on the hot decision-making task. Both the clinical group and the TD group performed only marginally above chance level. These results corroborate previous findings reporting no impairment on the HDT in children with ADHD relative to TD children (Geurts et al., 2006; Lambek et al., 2010; Masunami, Okazaki, & Maekawa, 2009). Since the preparation of this paper, a systematic review of hot EF in children and adolescents with ADHD reported mixed results regarding hot decision-making performance (Groen, Gaastra, Lewis-Evans, & Tucha, 2013). Of the six studies examining hot EF with different variants of the IGT, three studies found impaired hot decision-making in children and adolescents with ADHD relative to TD subjects (Garon, Moore, & Waschbusch, 2006; Hobson et al., 2011; Luman, Oosterlaan, Knol, & Sergeant, 2008). The three other studies reported no impairment in hot decision-making performance in children and adolescents with ADHD relative to TD counterparts (Geurts et al., 2006; Masunami et al., 2009; Toplak et al., 2005). Although all six studies examined decision-making with the IGT, different variants of the task were applied which might have influenced the outcome. Luman and colleagues (2008) used a variant of the IGT with three options (one advantageous, two disadvantageous) instead of four options (two advantageous, two disadvantageous). This adaptation involves a decrease in task complexity that might influence outcome. In less complex tasks, decision-making may to a greater extent be rationally guided because knowledge of the probability distribution is more likely. Decision-making deficits



reported in the study by Luman and colleagues (2008) may thus to a greater extent reflect cold EF impairment in subjects with ADHD than aberrant motivational processes (Groen et al., 2013). In the study by Garon and colleagues (2006), animal pictures were used instead of number symbols to indicate wins and losses on the IGT. The children were also rewarded with stickers, candy or toys in the study by Garon et al., (2006). By giving the participants tangible rewards instead of fictive rewards, there is the possibility that this would increase the motivational aspect in the IGT by evoking a stronger emotional response to gains and losses. Although Bowman and Turnbull (2003) found no effect on IGT performance in a study with adults whether the reward was symbolic or real money, emerging evidence indicates that the manipulation of the motivational component on a task will affect outcome in pediatric populations (Crone, Bullens, van der Plas, Kijkuit, & Zelazo, 2008; Figner, Mackinlay, Wilkening, & Weber, 2009). Interestingly, when manipulating the motivational salience in a gambling task, Crone and colleagues (2008) found that children and adolescents (8-18 years) made less risky choices when playing for another person than if they were playing for themselves. It is thus notable that two of the three studies reporting no decision-making deficits in children and adolescents with ADHD were applying the HDT version of the IGT, in which the subjects are asked to gain rewards (apples) for a donkey rather than for themselves (Geurts et al., 2006; Masunami et al., 2009). By comparison, another study using HDT not included in the review by Groen et al., (2013) found no abnormal decision-making performance in children with ADHD (Lambek et al., 2010). There is a possibility that the HDT used in our study represents a cooled down version of the IGT, being less sensitive to emotional/reward abnormalities than the standard IGT. Along these lines, the no-between group findings in our and other studies using the HDT (Geurts et al., 2006; Lambek et al., 2010; Masunami et al., 2009), may be a result of low construct validity of the task. In sum, inconsistent research findings on hot decision-making performance in children and adolescents with ADHD, may reflect that task complexity and motivational salience in the IGT version used are likely to have influenced the findings in previous studies.

Considering EF impairment in everyday life, we found as expected that children and adolescents with both ADHD subtypes were impaired on both hot and cold BRIEF scales. Consistent with our hypothesis, the ADHD-C group displayed more impairment on the cold Inhibition scale relative to the ADHD-I group, corroborating previous findings (Gioia et al., 2000; McCandless & O' Laughlin, 2007; Semrud-Clikeman, Walkowiak, Wilkinson, & Butcher, 2010). In addition, the ADHD-C group showed significantly more impaired scores on the cold Monitor scale, which is in line with findings by Gioia et al., (2000). Consistent with

previous findings, our results indicate that the cold Working Memory scale show best sensitivity in distinguishing ADHD children from TD children, while the cold Inhibit scale seems to be most effective in distinguishing between ADHD subtypes (McCandless & O'Laughlin, 2007). Further, our data provide evidence of some overlap between questions posed on the BRIEF and rating scales applied in the diagnostic process. In particular, the BRIEF Inhibition scale correlated strongly with hyperactivity/impulsivity items in the ADHD Rating Scale IV ( $r = .66$ ). This overlap between the BRIEF Inhibition scale and hyperactivity/impulsivity items formulated in the diagnostic process may thus result in some circularity when interpreting our findings. In sum, our findings corroborate previous studies showing that cold EF tests and BRIEF ratings reliably differentiate between children and adolescents with ADHD and TD children (Toplak et al., 2013).

In line with previous findings, correlation data displayed only modest associations between cold EF tests and BRIEF scales (V. A. Anderson et al., 2002; Toplak et al., 2009). Weak associations between neurocognitive test measures and BRIEF ratings raise the issues of whether these different measures tap the same or different underlying EF constructs. Since the preparation of this study, Toplak and colleagues (2013) examined this question in a review of 20 studies reporting correlations between cold EF tests and behavior ratings of EF. The majority of these studies (13) used the BRIEF to assess EF behavior. In general, associations between cold EF tests and behavior ratings of EF were small to modest (the median correlation reported was 0.19). Based on these findings, Toplak et al., (2013) argue that the two types of measures assess different aspects of EF. In agreement with McCue & Pramuka (1998), Toplak and colleagues (2013) suggest that cold EF tests are better interpreted as measures of maximal EF performance, which is in contrast to BRIEF ratings assessing typical EF performance (Toplak et al., 2013).

This distinction between typical and maximal performance can be aligned with the dual pathway model of ADHD that distinguishes between motivational and executive processes (Sonuga-Barke, 2003, 2005). Interestingly, correlation data showed no association between the HDT and cold EF tests either, which gives some differential validity to the hypothesis of separate hot and cold EF pathways in children and adolescents with ADHD (Sonuga-Barke, 2003). This separability of hot decision-making performance from cold EF tests is well documented in a review by Toplak and colleagues (2010). The majority of the 43 studies in that review reported no relationship between IGT performance and various cold EF tests. Our correlation data did not show any associations between BRIEF scales and the motivational driven HDT either. Based on the findings by Toplak and colleagues (2013), and our results

indicating weak associations between the EF tests (hot and cold) and BRIEF scales, there seems to be more separability than commonality among EF tests and behavior ratings of EF. Although different cold EF tests and BRIEF scales share some theoretical EF constructs, EF tests and BRIEF scales are not interchangeable aspects of EF.

Taken together, results from our first study show little support for different EF profiles in children and adolescents with ADHD-I and ADHD-C subtypes. However, as indicated in literature reviews by Groen et al., (2013) and Toplak et al., (2013), different psychometric characteristics of the EF measures applied to assess executive dysfunction in children and adolescents with ADHD may cause considerable variation in research findings.

In the second study, we therefore wanted to explore whether the lack of reliable EF differences between ADHD-C and the ADHD-I subtypes could be due to methodological aspects, such as a low test specificity in the neurocognitive tests applied. When investigating cognitive inhibition in a forced attention dichotic listening (DL) task, we found as expected impaired performance in both ADHD subgroups relative to TD children. In addition, the ADHD-I group demonstrated significantly greater impairment relative to the ADHD-C group. Our results corroborate previous findings reporting cognitive inhibition deficits in adults with ADHD using the DL procedure (Dramsdaahl, Westerhausen, Haavik, Hugdahl, & Plessen, 2011; Øie & Hugdahl, 2008). By comparison, no studies have reported impaired performance on the cognitive control condition in children and adolescents with ADHD using the DL test procedure (Davidson & Prior, 1978; Hiscock, Kinsbourne, Caplan, & Swanson, 1979; Prior, Sanson, Freethy, & Geffen, 1985; Øie, Rund, Sundet, & Bryhn, 1998). None of these studies explored the possibility of subtype differences potentially confounding previous findings. One possibility is that impairment was not detected when cognitive control deficits in the ADHD-I group is collapsed with the less impaired ADHD-C group. Our findings support the hypotheses set forward by Diamond (2005) that complex span tasks like the DL are more sensitive and are better at measuring the cognitive problems in children and adolescents with ADHD-I, and as far as we know this is the first study to demonstrate different degrees of inhibitory impairment between ADHD subtypes using the DL task

In sum, the main findings in the first and second studies seem to support and expand on previous studies reporting EF impairment in children and adolescent with ADHD. First, our results corroborate research findings reporting that cold EF deficits in general are equally impaired in children and adolescents with ADHD-C and ADHD-I subtypes. Second, our findings demonstrate that the DL task with improved test specificity shows differential diagnostic sensitivity between the ADHD-C and ADHD-I subtypes. This expands our

understanding of impaired EF in children and adolescents with ADHD subtypes. Finally, our finding that hot and cold EF seems to constitute separate EF pathways provides evidence for the multidimensionality of hot and cold EF in children and adolescents with ADHD.

### **5.1.2 Executive dysfunction and comorbidity in girls and boys with ADHD**

Research on gender differences has reported equally impaired cold EF in girls and boys with ADHD, but divergent comorbidity profiles. Internalizing symptoms seem to be more salient in females with ADHD, whereas externalizing problems are more widespread in male counterparts (Rucklidge, 2010; Seidman et al., 2005). Since less disruptive behavior in females with ADHD may cause underidentification and lack of treatment for this group, we explored in the third study whether there were gender differences in co-existing symptom severity and EF impairment in children and adolescents with ADHD. As expected, we found that cold EF tests had only a modest ability to categorize participants as ADHD or TD children. Further, there was only a slight difference in accuracy when categorizing participants as TD or ADHD in females versus males (79% versus 73% accuracy). These results corroborate previous findings reporting only a moderate ability of EF tests to classify children with ADHD. And the tests show no effectiveness in discriminating between genders (Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005; Rucklidge, 2010; Seidman et al., 2005; Willcutt et al., 2005). As expected, we found a significant effect of gender on parent-rated EF problems in everyday life. BRIEF ratings had a robust ability to categorize participants as ADHD or TD, and showed better discriminatory power in males than in females (96% versus 92% accuracy). The effect of gender on EF in everyday life in children and adolescents with ADHD remains somewhat unclear due to the scarcity of studies conducted on this topic. One study by Huizinga and colleagues (2011) reported comparable findings with elevated levels of executive problems in TD boys compared to TD girls assessed with the BRIEF. These results are comparable to data presented in the original version of the BRIEF, which show superior executive skills in girls compared to boys (Gioia et al., 2000). It is interesting to note that despite similar levels of neurocognitive problems in boys and girls with ADHD, EF impairment in everyday life seem to be more prominent in boys with ADHD, relative to female counterparts. This higher risk of executive behavior problems in males with ADHD may be one reason for the referral of more boys than girls for clinical evaluation of ADHD (Gaub & Carlson, 1997b). Our findings emphasise that overt behavior should not be the only basis for treatment, because this may cause underidentification of EF problems in girl with ADHD and hamper treatment for this group.

Moreover, co-existing symptom ratings were substantially better at distinguishing subjects with ADHD from TD in females (93% accuracy) than in males (86% accuracy). These findings indicate that co-existing psychological impairment may be even more reliably evident in females with ADHD compared to TD females than in males with ADHD relative to TD males (Rucklidge & Tannock, 2001). As expected, the most important co-existing symptom for categorizing females as ADHD or TD was increased self-reported internalizing problems, whereas parent-rated externalizing problems best categorized ADHD in males. These findings support and expand on previous studies reporting elevated levels of internalizing symptoms in girls with ADHD and more externalizing problems in boys with ADHD relative to TD counterparts (Gaub & Carlson, 1997b; Gershon, 2002; Quinn, 2008). Interestingly, our results indicate that girls with ADHD report more internalizing problems than their parents. As internalizing symptoms are less readily observed in children with ADHD, it is likely that parental reports of anxiety and depression in their children are less sensitive than self-reported internalizing problems (Grills & Ollendick, 2002; Rothen et al., 2009).

Taken together, there were three main findings in the third study. Despite similar levels of neurocognitive problems in boys and girls with ADHD, EF impairment in everyday life seem to be more prominent in boys with ADHD, relative to female counterparts. In contrast, co-existing problems seem to be more prevalent in girls with ADHD and in particular when based on the children's own ratings. One implication of these findings for clinical practice is that self-report scales may be used to obtain a better awareness of internalizing problems in females with ADHD.

### **5.1.3 Development of hot and cold EF in boys and girls with ADHD**

Previous studies have reported comparable developmental trajectories in cold EF for boys and girls with ADHD (Seidman, 2006). By comparison, longitudinal studies of hot EF in boys and girls with ADHD have been largely absent. In the fourth study, we were therefore interested in investigating the longitudinal course and impact of gender on hot and cold executive processes in children and adolescents with ADHD. As expected, we found that the ADHD group improved performance after two years, but remained impaired on all cold EF tests relative to TD children. Our findings are consistent with previous studies reporting persistent impairment in cold EF in children and adolescents with ADHD despite considerable maturation in childhood (Biederman et al., 2009; Seidman, 2006). Concurrently with this improvement in cold EF test performance, we observed a decline in co-existing internalizing and externalizing

problems and a decline in inattention and hyperactivity/impulsivity symptoms in the ADHD group. Although this decline in co-existing problems and ADHD symptoms may be due to the benefits of psychological and/or medical treatment, it is interesting to note that cold EF impairments seem to persist. Our results are consistent with findings in a 10-year longitudinal study of children with ADHD conducted by Biederman and colleagues (2009), which reported that cold EF impairments are relatively stable through childhood, and seem to be independent of the course of the disorder.

As expected, children and adolescents with ADHD did not improve on hot EF performance (decision-making) across time. Surprisingly, hot decision-making performance did not improve across time in the TD group either. These results are consistent with previous findings reporting that hot and cold EF mature at different rates in both TD children and children with ADHD (Crone & van der Molen, 2004; da Mata et al., 2011; Prencipe et al., 2011). A developmental arrest in hot EF performance relative to cold EF processes maps further onto theoretical models of hot and cold pathways suggesting a later and more gradual maturation of hot EF processes compared to cold EF processes (P. Anderson, 2002; Prencipe et al., 2011; Smith et al., 2012). Although we found no evidence of superior hot decision-making performance in boys compared to girls in general, our data provided evidence for a significant group x gender interaction. The females in the ADHD group displayed deteriorating decision-making performance across time, whereas females in the TD group displayed improved decision-making performance across time. Thus, when girls with ADHD outperformed TD counterparts at T1, it is notable that HDT (hot EF) results are comparable at T2. There is a growing consensus that TD boys outperform TD girls on decision-making tasks, but to the best of our knowledge no studies have investigated the effect of gender on decision-making in children and adolescents with ADHD (Crone et al., 2005; Garon & Moore, 2007; Overman, 2004). Nonetheless, our results with ADHD girls outperforming TD girls at T1, may potentially be caused by an effect where girls with ADHD follow decision-making strategies comparable to TD boys rather than TD females. In line with the findings by van den Bos and colleagues (2013), girls with ADHD may in general focus on overall gain similar to TD boys. In contrast, TD girls may to a greater extent focus on both win and loss frequencies as well as overall gain. The more detailed oriented decision-making strategy may cause TD girls to need more trials to reach the same level of performance as girls with ADHD at T1.

Interestingly, a decline in decision-making performance across time in girls with ADHD may indicate a developmental delay, or even a regression in hot EF processes as previously reported by Anderson (2002) and Smith and colleagues (2012). When investigating

decision-making with the IGT in TD children and adolescents from 8 to 17 years, Smith, Xiao and Bechara (2012) found a curvilinear effect of age. The youngest participants failed to develop a preference for any of the doors, but outperformed the more reward-driven participants in early adolescence (10-13 years). Decision-making improved again from 14 years and peaked at the age of 17. A potential interpretation in line with the findings reported by Smith and colleagues (2012) may be that girls with ADHD outperformed TD girls at T1 because they were less cognitively developed, and, as such, less sensitive to the high reward doors (causing overall loss). With increasing age, however, decision-making performance declines in girls with ADHD as activation in the nucleus accumbens reward system increases, and they become more sensitive to the large payoffs on the hot EF task (Ernst et al., 2005; Galvan et al., 2006). A curvilinear effect of age on decision-making performance may explain why previous findings have found decision-making impairment to be more evident in adolescents with ADHD than in children with ADHD (Geurts et al., 2006; Lambek et al., 2010; Masunami et al., 2009; Toplak et al., 2005).

It has been suggested that performance on decision-making tasks may be explained by cold EF abilities such as working memory and cognitive inhibition (Dunn, Dalgleish, & Lawrence, 2006). Decision-making performance could thus be mediated by the maturation of cold EF abilities. However, as reported above, both the TD and ADHD children displayed improved cold EF after two years. If hot EF performance was mediated by cold EF abilities, improved hot EF performance should have been demonstrated in both the ADHD and TD subjects, which is clearly not the case. Further, in a review of 43 studies examining the association between IGT/HDT and other cognitive abilities, Toplak, Sorge, Benoit, West, and Stanovich (2010) found no strong associations between IGT/HDT performance and working memory, cognitive inhibition, cognitive flexibility, or intelligence. Clearly the role of gender in ADHD decision-making warrants further study. However it seems plausible that hot EF processes show a protracted period of development, which can be distinguished from a more linear development in cold EF processes.

In sum, the main findings from the fourth study corroborate and expand on previous studies reporting separate developmental trajectories in hot and cold EF across development in children and adolescents with ADHD. As expected, we found that cold EF impairments persist, despite a considerable improvement across time in boys and girls with ADHD. Second, the persistence of cold EF impairments despite a decline in symptoms indicates that cold executive dysfunctions are mainly trait dependent rather than state dependent phenotypes in children and adolescents with ADHD. Finally, deteriorating decision-making performance across time in

girls with ADHD and improved performance in TD girls indicates divergent developmental trajectories of hot EF in girls with ADHD relative to TD counterparts.

## **5.2 Methodological issues**

### **5.2.1 Representativeness and generalisability of results**

The clinical sample in this study was recruited as consecutive referrals for assessment of ADHD from all seven child and adolescent outpatient mental health centers in two neighboring Norwegian counties (Hedmark and Oppland). The study population included in this study is therefore considered to have a relatively high degree of representativity for individuals with ADHD receiving treatment at mental health units in Norway. However, the sample was drawn from a clinical population and represents those who are willing to seek help in a mental health clinic. Our results and conclusions may therefore not be generalizable to the entire ADHD population.

The participation rejection rate was unfortunately not registered. Consequently, it remains unclear whether those rejecting participation were lower functioning than those included in the study. There is also a possibility that children with parents having an academic interest in the research project were more likely to participate, potentially biasing the inclusion towards more “high functioning” individuals. A higher educational level in the mothers of TD children compared to other studies in the field may have resulted in a more high functioning participants in our sample. However, the educational level of the mothers in the TD group was nearly equal to that of TD mothers in other studies conducted in Norway (Heiervang, Mednick, Sundet, & Rund, 2010; Øie, Sundet, & Ueland, 2011).

### **5.2.2 Possible confounding factors**

Learning disabilities, prior stimulant medication and other medical problems affecting the central nervous system (i.e. prematurity < 36 weeks) were specified as a priori exclusion criteria in order to control for potential confounders known to influence neurocognition. However, it is difficult to control for all potential confounding variables in a naturalistic study. When demographic and clinical variables known to influence neurocognition differed between groups, these potential confounders were statistically controlled for when possible.

General cognitive functioning (IQ) is an important confounding factor when assessing EF, as these variables are highly correlated (Ardila, Pineda, & Rosselli, 2000; Lezak, Howieson, & Loring, 2004). IQ below 70 was controlled for by the study design, and the influence of IQ on EF outcome was investigated by entering it as a covariate in the first and



fourth studies. The main findings in these two studies remained significant when controlling for IQ, suggesting that IQ did not confound the main results. IQ was not controlled for in the second study. The main findings in the third study were based on non-parametric statistical analyses. There is, however, a continuing debate in the literature as to whether EF data should be corrected for overall IQ, because controlling for IQ removes a portion of the variance that is related to ADHD (Denckla, 1996; Nigg, 2001).

The beneficial effect of stimulant medication on neurocognition and behavioral problems in children and adolescents with ADHD is well documented (Abikoff et al., 2004; Swanson, Baler, & Volkow, 2011; van der Oord, Prins, Oosterlaan, & Emmelkamp, 2008). As no participants with ADHD were prescribed stimulants or had previously been medicated with stimulants at baseline we can rule out this as a confounding variable at T1. At the time of reassessment, however, 34.4% of the ADHD participants were taking stimulant medication ( $n = 42$ ; Equasym, Concerta and Ritalin). In order to control for the potential impact of stimulant medication on EF outcome at T2, the children taking stimulants discontinued taking their medication at least 24 hours before re-assessment at T2 (Coghill et al., 2013). The clinical effect of stimulants lasts typically 3-12 hours, and it is generally assumed that the neurocognitive impairment returns immediately when the drug effect diminishes (Huang, Wang, & Chen, 2012). In addition, all analyses in the fourth study were repeated excluding the participants on stimulant medication. The exclusion of participants taking stimulants between T1 and T2 did not have any impact on the main findings. Thus, we believe that EF outcome assessed with EF tests are reasonably well controlled for at T2.

With increasing age, neurocognitive performance will improve in both the TD group and the ADHD group. The effect of age will certainly have an impact on results in our studies. Mean age in the ADHD group and TD group did not differ between groups at baseline or at follow-up. Based on this, age was not believed to confound between group findings. However, the relatively large age span in our population (8-17 years at T1) may cause some within group variability in test performance, since EF is known to undergo considerable change throughout childhood and adolescence (P. Anderson, 2002; Best & Miller, 2010). In the second study, the TD participants were randomly drawn to match the age of the ADHD participants. One possibility could have been to match the groups on age in the other three studies as well, but this was not done due to the limited number of participants in the different groups. In order to overcome the possibility of within group variability on EF test measures at T2 and obtain greater age homogeneity, the age span in the fourth study was restricted to 9-16 years at T1 instead of 8-17 years.

### **5.2.3 Attrition**

The fourth study had a longitudinal design with two measurement points across two years. Longitudinal studies make it possible to observe developmental trajectories of different EF components, instead of drawing interpretations out of cross-sectional observations. One of the challenges in longitudinal studies, however, is selective attrition that may bias results. Therefore, the identification of possible differences in age, gender and other relevant confounders between responders and non-responders is of crucial importance. In general, there were few drop outs (4%) from T1 to T2. Age and IQ did not differ in ADHD participants available for re-assessment and those three lost to follow-up, which suggests that there was no selection bias in these variables from T1 to T2. Further, participants with ADHD and TD children did not differ significantly with regard to gender distribution at T1 or T2. When taking into consideration the low drop-out rate and the insignificant impact of age, IQ and gender bias on results at T2, it seems plausible that attrition is not a major concern in the fourth study.

### **5.2.4 Measurement**

EF and emotional problems in children and adolescents with ADHD was the focus in the current study. In order to conduct a quantitative investigation of these two concepts, it is necessary to operationalize the meaning of EF and emotional problems in order to make these two concepts available for measurement. Consequently, the quality of measurement methods included in this study are of crucial importance for the validity of our research findings. Only standardised neurocognitive tests and rating scales demonstrating acceptable psychometric properties were used. Further, clinicians were trained in neurocognitive testing and diagnostic assessment prior to inclusion of participants. Frequent supervisory meetings were also held in order to ensure reliable assessments between clinicians.

The neurocognitive tests used in this study overlap with many previous studies and are widely accepted. Measures included from the Delis-Kaplan Executive Function System (D-KEFS) and WISC-IV to assess cold EF have all been tested for good psychometric properties (Delis et al., 2001; Groth-Marnat, 2003; Wechsler, 2004; Wodka et al., 2008). Nevertheless, low test specificity and a lack of ecological validity have been raised as a concern regarding neurocognitive assessment. EF tests are typically highly complex and require other skills than those targeted for assessment (Miyake et al., 2000). As a result, non-EF components like reading problems, language deficits, and reduced cognitive tempo may influence the outcome

on EF tests. As illustrated in our second study, test specificity was improved when using a test paradigm (DL) that is simple to understand and not dependent on rapid cognitive tempo and reading skills.

The structured setting of a test room will typically not reflect the natural surroundings of subjects being tested. As a result, cold EF test results attained in a test room will not reflect EF in the real world (Sbordone, 2000). Furthermore, the lack of motivation in a test situation may also influence performance (Sonuga-Barke, 2005). Concerns regarding the ecological validity of EF assessment were dealt with by including behavior ratings of EF in everyday life (BRIEF) and a task assessing EF under more emotion-laden circumstances (HDT).

As previously discussed, the use of different variants of the IGT may have caused inconsistent research findings when exploring hot decision-making in children and adolescents with ADHD (Groen et al., 2013). This raises the concern as to what degree IGT and HDT actually measure the construct hot EF. Or to be more specific, the interaction between motivational processes and cold EF. When examining the construct validity of the IGT, Buelow and Suhr (2009) found that emotional processing was reliably associated with hot decision-making performance. This supports the validity of using IGT and HDT as a hot EF measure. The authors stress, however, the lack of studies examining the ecological validity of IGT. In our first study, correlations between the HDT and BRIEF measures revealed no associations between these two measures aimed at assessing EF in “real-world” situations. Although the BRIEF is not developed to assess hot decision-making, our results indicate that the ecological validity of IGT/HDT needs further clarification.

There is also the possibility that no between-group findings on the HDT in our first study was caused by a so called floor effect. In a study by Lambek and colleagues (2010), only half of the participants from 7-14 years (ADHD and TD) could give an accurate account of the rationale behind the HDT after completing the task. By comparison, since the participants were relatively young (8-17 years) in our first study as well, the contingency schedules involved in the HDT may have been too difficult to understand and solve for the children. If this was the case, this would have compromised the HDT results at baseline.

The assessments of emotional problems were based on widely accepted rating scales demonstrating good psychometric properties. We consider the use of both self-report and parent rating scales as an advantage. Child reports may be more informative regarding internalizing problems, whereas parent ratings seem to be more valuable concerning externalizing problems in paediatric populations (Goodman, Ford, Richards, Gatward, & Meltzer, 2000).

### **5.3 Implications**

The findings in this study may have implications for our understanding of the neurocognitive and emotional problems in children and adolescents with ADHD. First, our results will be discussed in the context of the dual-pathway model of ADHD. Second, the clinical implication of our results will be outlined.

#### **5.3.1 Theoretical implications**

As mentioned in the introduction of this thesis, there is considerable consensus regarding executive dysfunction as a central source of the disability associated with ADHD (Willcutt et al., 2005). However, recent theories argue that there may be multiple pathways explaining ADHD symptoms (Sonuga-Barke, 2005). According to Sonuga-Barke's (2005) dual pathway model of ADHD, executive dysfunction (cold EF) and motivational dysfunction (hot EF) constitute two separate pathways for ADHD related problems.

The first study showed impaired cold EF in subjects with ADHD, but no between group differences in hot EF performance. This finding provides evidence of the separability of hot and cold EF pathways. Likewise, the lack of correlations found between hot EF performance and cold EF tests give further support to the hypothesis of separate pathways in ADHD (Sonuga-Barke, 2005). The separability of hot EF performance from cold EF tests is well documented, but the degree of impairment in hot EF processes in children and adolescents with ADHD need further clarification (Groen et al., 2013; Toplak et al., 2010). The findings from the fourth study revealed deteriorating hot EF performance across time in girls with ADHD, but improved cold EF performance in boys and girls with ADHD. This finding seems to indicate separate developmental trajectories in hot and cold EF pathways. Thus, results from the first and fourth studies seem to support a dual pathway model involving divergent hot and cold EF trajectories in children and adolescents with ADHD.

#### **5.3.2 Clinical implications**

The findings from this thesis may have implications for treatment as well as for neurocognitive assessment in children and adolescents with ADHD.

Results from the first study corroborated previous studies showing that impaired cold EF is a central deficit in children and adolescents with ADHD. Consequently, neurocognitive assessment is important in daily clinical practice to guide treatment and intervention plans. Further, our results demonstrate better sensitivity and specificity for behavior ratings of EF

than for neurocognitive tests. These findings indicate that the BRIEF may be a valuable tool to guide treatment, and more suitable than neurocognitive testing to differentiate between ADHD-C and ADHD-I.

In the second study, results show that ADHD subtypes differ in degree of cognitive inhibition impairment. By using an experimental approach that is simple to understand and not dependent on rapid cognitive tempo or reading skills, children with the ADHD-I subtype demonstrated more problems with cognitive inhibition relative to children with the ADHD-C subtype. Due to the fact that reading problems, language deficits, and reduced cognitive tempo are more common in ADHD-I than in ADHD-C, our results illustrate the importance of not using tests or methods that involve multiple executive processes when examining possible EF differences between clinical subtypes. Considerable problems inhibiting irrelevant auditory syllables in the ADHD-I group may be associated with difficulties in ignoring irrelevant and interfering “noise” in the classroom or in social settings. Previous research has documented that many individuals with ADHD get lower grades in school, where the verbal component is tremendously important (Faraone, Biederman, Monuteaux, Doyle, & Seidman, 2001). An experimentally-based study like the current one, which shows impairments in cognitive inhibition in a highly structured environment, testifies to the importance of addressing the issue of techniques to improve cognitive inhibition for this group of young students.

Our findings in the third study emphasize the combination of self-report and parent rating scales for the identification of different comorbid symptom expression in boys and girls diagnosed with ADHD. We found that parent ratings were more informative regarding behavioral problems in boys with ADHD, whereas self-report scales may increase awareness of internalizing problems particularly salient in females with ADHD. Given the almost universal phenomenon of “comorbidity” in ADHD, our findings emphasize the combined value of self-report and parent rating scales for the identification of comorbid symptoms in boys and girls already diagnosed with ADHD.

In the fourth study, we found that cold EF deficits persist despite a decline in ADHD symptoms and comorbidity. This finding adds to a growing literature indicating that cold executive dysfunction is a stable endophenotypic trait in ADHD (Biederman et al., 2009). Clearly, as cold EF deficits are strongly associated with poor behavioral, social, educational and occupational outcomes in ADHD populations, the impairment level needs to be considered when ADHD symptoms are remitting as well (Barkley & Murphy, 2010; Biederman et al., 2004; Biederman, Petty, et al., 2006).

#### **5.4 Strengths, limitations and future research**

The inclusion of a relatively big sample size compared to other studies on hot EF, and a longitudinal design with few drop outs (4%) constitutes major strengths of this study. The interaction between EF and motivational processes in children and adolescents with ADHD remains unclear, and in particular the longitudinal course of hot EF is largely unexplored. The cross-sectional and longitudinal assessment of both hot and cold EF have extended the investigation of EF deficits in line with newer theoretical models of ADHD. The assessment of EF with BRIEF has provided a more nuanced picture of EF in children and adolescents with ADHD, as behavior ratings of EF demonstrate better sensitivity to EF problems than neurocognitive tests. Additional strengths are the use of both self-report and parent ratings when assessing co-existing symptoms in a well-balanced sample of boys and girls with ADHD. Finally, the inclusion of ADHD participants never medicated with stimulants prior to assessment at T1 constitutes another major strength of this study. As stimulant medication may have a beneficial effect on neurocognition, our test results from stimulant-naïve subjects represent a valid picture of EF deficits in children and adolescents with ADHD. In addition, all of the ADHD participants, except one, were tested without stimulant medication at T2.

One potential limitation is that the inclusion of only one hot EF task may have been somewhat unbalanced relative to the number of cold EF tests in this study. This concern seems plausible in light of a recent literature review reporting considerable psychometric variability in different hot EF tasks (Groen et al., 2013). The large age span (9-16 years at T1) might represent another limitation. As EF in children and adolescents develop and change rapidly throughout childhood, large age spans may cause large within group variability. Heterogeneous EF results in the ADHD and TD groups may cause low discriminating power between groups and result in type II errors.

Based on the findings in this thesis, future studies should include DL test measures, since this paradigm has the power to detect subtle differences in EF not detected by conventional test procedures. The DL paradigm could be used in further explorations of EF subcomponents and the developmental trajectories of these EF subcomponents in children and adolescents with ADHD. Moreover, future research should include different hot EF test measures when exploring the development of these skills in children and adolescents with ADHD. In particular, the distinction between implicit and explicit hot EF tasks will be crucial to consider. Implicit hot EF tasks reflect to a greater extent emotional and affective responses to the options of choice, whereas responses in explicit hot EF tasks are more determined by

cognitive and rational processes (Groen et al., 2013). Clearly, test results from implicit or explicit hot EF test measures will reflect different cognitive skills.

## **5.5 Ethics**

The study was approved by the Regional Committee for Medical Research Ethics in Eastern Norway (REK-Øst) and by the Privacy protection ombudsman for research at Innlandet Hospital Trust. It was conducted in accordance with the Helsinki Declaration of the World Medical Association Assembly. In order to ensure a free informed consent, all candidates for the study were duly informed about the research project in an information letter and in the consent form. All parents/caregivers and participants above 12 years gave written informed consent in accordance with the Research Ethics Committee in Eastern Norway. Children under the age of 12 years provided oral consent to participate. The hospital's research server was used for data storage and only authorized persons had access to the archives securing the confidentiality of participants. The comprehensive assessment at T1 with neurocognitive testing and diagnostic interviews was clinically relevant and in the interest of participants. All participants and their parents were offered information about the neurocognitive test results and the diagnostic judgements. However, reassessment at T2, which was relatively time consuming, may have challenged the motivation of some parents and participants. After 2 years, the clinical relevance of the comprehensive assessment was less salient, and to some extent primarily in the interest of the researchers. It is therefore always important to keep in mind that the disadvantages for the participants enrolled in a research project should be avoided or kept at a minimum.

## **6. CONCLUSION**

In sum, few EF measures were shown to differentiate between ADHD subtypes; nor were any relationships revealed between the hot decision-making task and the other EF measures. The BRIEF seems more suitable than neurocognitive testing to differentiate between the subtypes ADHD-C and ADHD-I. The dichotic listening task did, however, demonstrate a significantly higher degree of inhibitory impairment in ADHD-I compared to ADHD-C in a highly structured environment. These findings suggest that psychometric characteristics of the EF measures applied may blur the identification of distinct EF profiles in ADHD-I and ADHD-C subtypes. Further, females with ADHD are more accurately distinguished from TD children by the presence of co-existing symptoms, particularly physiological components of anxiety. On the other hand, parental reports of EF impairments are substantially better at distinguishing

subjects with ADHD from TD males. Our findings emphasize the combination of self-report and parent rating scales for the identification of different comorbid symptom expression in boys and girls diagnosed with ADHD. Finally, cold EF impairments persist across time, despite a considerable improvement after 2 years in boys and girls with ADHD. In contrast, girls with ADHD show deteriorating hot EF performance across time. Our findings indicate that hot and cold EF show divergent developmental trajectories, which are relatively independent from co-existing problems and ADHD symptomatology in children and adolescents with ADHD. Future research should investigate different developmental trajectories of hot and cold EF in a longitudinal study, as most findings to date are based on cross-sectional comparisons.



## 7. REFERENCES

- Abikoff, H., Hechtman, L., Klein, R. G., Weiss, G., Fleiss, K., Etcovitch, J., . . . Pollack, S. (2004). Symptomatic improvement in children with ADHD treated with long-term methylphenidate and multimodal psychosocial treatment. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(7), 802-811.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.
- Adams, Z. W., Derefinco, K. J., Milich, R., & Fillmore, M. T. (2008). Inhibitory functioning across ADHD subtypes: recent findings, clinical implications, and future directions. *Developmental Disabilities Research Reviews, 14*(4), 268-275.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed.). Washington, DC: American Psychiatric Association.
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71-82.
- 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.
- Angold, A., Costello, E. J. , Messer, S. C., Pickles, A., Winder, F. , & Silver, D. (1995). The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research, 5*, 1-12.
- Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology, 15*(1), 31-36.
- Arnold, L. E. (1996). Sex differences in ADHD: conference summary. *Journal of Abnormal Child Psychology, 24*(5), 555-569.
- Bakker, D. J., van der Vlugt, H., & Claushuis, M. (1978). The reliability of dichotic ear asymmetry in normal children. *Neuropsychologia, 16*(6), 753-757.
- Banerjee, T. D., Middleton, F., & Faraone, S. V. (2007). Environmental risk factors for attention-deficit hyperactivity disorder. *Acta Paediatrica, 96*(9), 1269-1274.
- 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*: The Guilford Press.

- Barkley, R. A., & Murphy, K. R. (2010). Impairment in occupational functioning and adult ADHD: the predictive utility of executive function (EF) ratings versus EF tests. *Archives of Clinical Neuropsychology, 25*(3), 157-173.
- Barkley, R. A., Murphy, K. R., & Fischer, M. (2010). *ADHD in adults: What the science says*: The Guilford Press.
- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1-3), 7-15.
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641-1660.
- Biederman, J. (2005). Attention-deficit/hyperactivity disorder: a selective overview. *Biological Psychiatry, 57*(11), 1215-1220.
- Biederman, J., Kwon, A., Aleardi, M., Chouinard, V. A., Marino, T., Cole, H., . . . Faraone, S. V. (2005). Absence of gender effects on attention deficit hyperactivity disorder: findings in nonreferred subjects. *American Journal of Psychiatry, 162*(6), 1083-1089.
- Biederman, J., Mick, E., Faraone, S. V., Braaten, E., Doyle, A., Spencer, T., . . . Johnson, M. A. (2002). Influence of gender on attention deficit hyperactivity disorder in children referred to a psychiatric clinic. *American Journal of Psychiatry, 159*(1), 36-42.
- Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., . . . Faraone, S. V. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology, 72*(5), 757-766.
- Biederman, J., Monuteaux, M. C., Mick, E., Spencer, T., Wilens, T. E., Silva, J. M., . . . Faraone, S. V. (2006). Young adult outcome of attention deficit hyperactivity disorder: a controlled 10-year follow-up study. *Psychological Medicine, 36*(2), 167-179.
- 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.
- Biederman, J., Petty, C. R., Ball, S. W., Fried, R., Doyle, A. E., Cohen, D., . . . Faraone, S. V. (2009). Are cognitive deficits in attention deficit/hyperactivity disorder related to the course of the disorder? A prospective controlled follow-up study of grown up boys with persistent and remitting course. *Psychiatry Research, 170*(2-3), 177-182.

- Bodnar, L. E., Prahme, M. C., Cutting, L. E., Denckla, M. B., & Mahone, E. M. (2007). Construct validity of parent ratings of inhibitory control. *Child Neuropsychology*, *13*(4), 345-362.
- Boehnke, K., Sillbereisen, R. K., Reynolds, C. R., & Richmond, B. O. (1986). What I think and feel: German experience with the revised form of the children's manifest anxiety scale. *Personality and Individual Differences*, *7*(4), 553-560.
- Bowman, C. H., & Turnbull, O. H. (2003). Real versus facsimile reinforcers on the Iowa Gambling Task. *Brain and Cognition*, *53*(2), 207-210.
- Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.
- Breiman, L., & Cutler, A. (2003). Random forests. Version 4.0. <http://www.stat.berkeley.edu/users/breiman/RandomForests/>.
- Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa Gambling Task. *Neuropsychology Review*, *19*(1), 102-114.
- Castellanos, F. X., Sonuga-Barke, E. J., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*(3), 117-123.
- Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201-216.
- Chhabildas, N., Pennington, B. F., & Willcutt, E. G. (2001). A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *Journal of Abnormal Child Psychology*, *29*(6), 529-540.
- Coghill, D., Seth, S., Pedroso, S., Usala, T., Currie, J., & Gagliano, A. (2013). Effects of Methylphenidate on Cognitive Functions in Children and Adolescents with Attention-Deficit/Hyperactivity Disorder: Evidence from a Systematic Review and a Meta-Analysis. *Biological Psychiatry*. doi: 10.1016/j.biopsych.2013.10.005
- Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: checklists, screens, and nets. *Journal of the American Academy of Child & Adolescent Psychiatry*, *27*(6), 726-737.
- Costello, E. J., Benjamin, R., Angold, A., & Silver, D. (1991). Mood variability in adolescents: A study of depressed, nondepressed and comorbid patients. *Journal of Affective Disorders*, *23*(4), 199-212.

- Crone, E. A., Bullens, L., van der Plas, E. A., Kijkuit, E. J., & Zelazo, P. D. (2008). Developmental changes and individual differences in risk and perspective taking in adolescence. *Development and Psychopathology*, *20*(4), 1213-1229.
- Crone, E. A., Bunge, S. A., Latenstein, H., & van der Molen, M. W. (2005). Characterization of children's decision making: sensitivity to punishment frequency, not task complexity. *Child Neuropsychology*, *11*(3), 245-263.
- Crone, E. A., & van der Molen, M. W. (2004). Developmental changes in real life decision making: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology*, *25*(3), 251-279.
- da Mata, F. G. , Silva, N. F., Menezes, L. G. , Paiva de Moraes, P. H. , Mattos, P. , Fuentes, D. , . . . Malloy-Diniz, L. (2011). Neuropsychological assessment of the decision making process in children and adolescents: an integrative review of the literature. *Revista de Psiquitria Clínica*, *38*(3), 106-115.
- Davidson, E. M., & Prior, M. R. (1978). Laterality and selective attention in hyperactive children. *Journal of Abnormal Child Psychology*, *6*(4), 475-481.
- Delis, D. , Kaplan, E., & Kramer, J. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. Norwegian version. Stockholm: Pearson Assessment.
- Denckla, M. B. (1996). Biological correlates of learning and attention: what is relevant to learning disability and attention-deficit hyperactivity disorder? *Journal of Developmental and Behavioral Pediatrics*, *17*(2), 114-119.
- Diamond, A. (2005). Attention-deficit disorder (attention-deficit/ hyperactivity disorder without hyperactivity): a neurobiologically and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Development and Psychopathology*, *17*(3), 807-825.
- Dramsdaahl, M., Westerhausen, R., Haavik, J., Hugdahl, K., & Plessen, K. J. (2011). Cognitive control in adults with attention-deficit/hyperactivity disorder. *Psychiatry Research*, *188*(3), 406-410.
- Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews*, *30*(2), 239-271.
- DuPaul, G. J., Power, T. J. , Anastoupolous, A. D. , & Reid, R. (1998). ADHD rating scale – IV. Checklists, norms & clinical interpretation. New York: Guilford.
- Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., . . . Pine, D. S. (2005). Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage*, *25*(4), 1279-1291.

- Fallmyr, O., & Egeland, J. (2011). Psychometric properties of the Norwegian version of BRIEF – for children from 5 to 18 years old. *Journal of the Norwegian Psychological Association, 48*, 339-343.
- Faraone, S. V., Biederman, J., Monuteaux, M. C., Doyle, A. E., & Seidman, L. J. (2001). A psychometric measure of learning disability predicts educational failure four years later in boys with Attention-Deficit/Hyperactivity Disorder. *Journal of Attention Disorders, 4*(4), 220-230.
- Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A., & Sklar, P. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57*(11), 1313-1323.
- Ferrando, P. J. (1994). Factorial structure of the revised children manifest anxiety scale in a Spanish sample: Relations with Eysenck personality dimensions. *Personality and Individual Differences, 16*(5), 693-699.
- Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology, 35*(3), 709-730.
- Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience, 26*(25), 6885-6892.
- Garon, N., & Moore, C. (2007). Developmental and gender differences in future-oriented decision-making during the preschool period. *Child Neuropsychology, 13*(1), 46-63.
- Garon, N., Moore, C., & Waschbusch, D. A. (2006). Decision making in children with ADHD only, ADHD-anxious/depressed, and control children using a child version of the Iowa Gambling Task. *Journal of Attention Disorders, 9*(4), 607-619.
- Gaub, M., & Carlson, C. L. (1997a). Behavioral characteristics of DSM-IV ADHD subtypes in a school-based population. *Journal of Abnormal Child Psychology, 25*(2), 103-111.
- Gaub, M., & Carlson, C. L. (1997b). Gender differences in ADHD: a meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(8), 1036-1045.
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders, 5*(3), 143-154.
- Geurts, H. M., van der Oord, S., & Crone, E. A. (2006). Hot and cool aspects of cognitive control in children with ADHD: decision-making and inhibition. *Journal of Abnormal Child Psychology, 34*(6), 813-824.

- Giedd, J. N., Lalonde, F. M., Celano, M. J., White, S. L., Wallace, G. L., Lee, N. R., & Lenroot, R. K. (2009). Anatomical brain magnetic resonance imaging of typically developing children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(5), 465-470.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology*, 6(3), 235-238.
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41, 645-656.
- Grills, A. E., & Ollendick, T. H. (2002). Issues in parent-child agreement: The case of structured diagnostic interviews. *Clinical Child and Family Psychology Review*, 5(1), 57-83.
- Groen, Y., Gaastra, G. F., Lewis-Evans, B., & Tucha, O. (2013). Risky behavior in gambling tasks in individuals with ADHD - a systematic literature review. *PLoS One*, 8(9), e74909.
- Groth-Marnat, G. (2003). *Handbook of psychological assessment*. Hoboken, NJ: John Wiley & Sons.
- Harper, L. V., & Kraft, R. H. (1986). Lateralization of receptive language in preschoolers: Test-retest reliability in a dichotic listening task. *Developmental Psychology*, 22(4), 553-556.
- Heiervang, K. S., Mednick, S., Sundet, K., & Rund, B. R. (2010). The Chernobyl accident and cognitive functioning: a study of Norwegian adolescents exposed in utero. *Developmental Neuropsychology*, 35(6), 643-655.
- Hinshaw, S. P., Owens, E. B., Sami, N., & Fargeon, S. (2006). Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *Journal of Consulting and Clinical Psychology*, 74(3), 489-499.
- Hiscock, M., Kinsbourne, M., Caplan, B., & Swanson, J. M. (1979). Auditory attention in hyperactive children: Effects of stimulant medication on dichotic listening performance. *Journal of Abnormal Psychology*, 88(1), 27-32.
- Hobson, C. W., Scott, S., & Rubia, K. (2011). Investigation of cool and hot executive function in ODD/CD independently of ADHD. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 52(10), 1035-1043.

- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65-70.
- Huang, Y. S., Wang, L. J., & Chen, C. K. (2012). Long-term neurocognitive effects of methylphenidate in patients with attention deficit hyperactivity disorder, even at drug-free status. *BMC Psychiatry*, 12(1), 1-7.
- Hugdahl, K., & Andersson, L. (1986). The “forced-attention paradigm” in dichotic listening to CV-syllables: A comparison between adults and children. *Cortex*, 22(3), 417-432.
- Huizinga, M., & Smidts, D. P. (2011). Age-related changes in executive function: A normative study with the Dutch version of the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychology*, 17(1), 51-66.
- Ivanova, M. Y., Dobrean, A., Dopfner, M., Erol, N., Fombonne, E., Fonseca, A. C., . . . Chen, W. J. (2007). Testing the 8-syndrome structure of the child behavior checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, 36(3), 405-417.
- Jonsdottir, S., Bouma, A., Sergeant, J. A., & Scherder, E. J. (2006). Relationships between neuropsychological measures of executive function and behavioral measures of ADHD symptoms and comorbid behavior. *Archives of Clinical Neuropsychology*, 21(5), 383-394.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(7), 980-988.
- Kopp, S., Kelly, K.B., & Gillberg, C. (2010). Girls with social and/or attention deficits: a descriptive study of 100 clinic attenders. *Journal of Attention Disorders*, 14(2), 167.
- Kovacs, M. (1983). *The Children's Depression Inventory: A self-rated depression scale for school-aged youngsters*. Pittsburgh, PA: University of Pittsburgh School of Medicine.
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1-26.
- Lambek, R., Tannock, R., Dalsgaard, S., Trillingsgaard, A., Damm, D., & Thomsen, P. H. (2010). Validating neuropsychological subtypes of ADHD: how do children with and without an executive function deficit differ? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51(8), 895-904.
- Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., . . . Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, 36(4), 1065-1073.

- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment* (4 ed.). New York: Oxford University Press.
- Luman, M., Oosterlaan, J., Knol, D. L., & Sergeant, J. A. (2008). Decision-making in ADHD: sensitive to frequency but blind to the magnitude of penalty? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(7), 712-722.
- Luman, M., Oosterlaan, J., & Sergeant, J. A. (2005). The impact of reinforcement contingencies on AD/HD: a review and theoretical appraisal. *Clinical Psychology Review*, *25*(2), 183-213.
- Lundervold, A. J., Breivik, K., Posserud, M. B., Stormark, K. M., & Hysing, M. (2013). Symptoms of depression as reported by Norwegian adolescents on the Short Mood and Feelings Questionnaire. *Frontiers in Psychology*, *4*, 613. doi: 10.3389/fpsyg.2013.00613
- Masunami, T., Okazaki, S., & Maekawa, H. (2009). Decision-making patterns and sensitivity to reward and punishment in children with attention-deficit hyperactivity disorder. *International Journal of Psychophysiology*, *72*(3), 283-288.
- 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.
- McCue, M., & Pramuka, M. (1998). Functional Assessment. In G Goldstein & S Beers (Eds.), *Rehabilitation* (pp. 113-129). New York: Plenum.
- Milich, R., Balentine, A. C., & Lynam, D. R. (2001). ADHD combined type and ADHD predominantly inattentive type are distinct and unrelated disorders. *Clinical Psychology*, *8*, 463-488.
- Miller, M., Sheridan, M., Cardoos, S. L., & Hinshaw, S. P. (2013). Impaired decision-making as a young adult outcome of girls diagnosed with attention-deficit/hyperactivity disorder in childhood. *Journal of the International Neuropsychological Society*, *19*(1), 110-114.
- Millichap, J. G. (2008). Etiologic classification of attention-deficit/hyperactivity disorder. *Pediatrics*, *121*(2), e358-365.
- Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8-14.
- 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.
- Nigg, J. T. (2001). Is ADHD a disinhibitory disorder? *Psychological Bulletin*, 127(5), 571-598.
- Nigg, J. T. (2005). Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: the state of the field and salient challenges for the coming decade. *Biological Psychiatry*, 57(11), 1424-1435.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, 57(11), 1224-1230.
- Nøvik, T. S. (1999). Validity of the Child Behaviour Checklist in a Norwegian sample. *European Child & Adolescent Psychiatry*, 8(4), 247-254.
- Nøvik, T. S. (2000). Child Behavior Checklist item scores in Norwegian children. *European Child & Adolescent Psychiatry*, 9(1), 54-60.
- O'Hare, E. D., Sowell, E. R. , Nelson, C. A. , & Luciana, M. (2008). Imaging developmental changes in gray and white matter in the human brain *Handbook of developmental cognitive neuroscience* (pp. 23-38). Cambridge, MA: MIT Press.
- Overman, W. H. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain and Cognition*, 55(1), 134-147.
- Pela, O. A., & Reynolds, C. R. (1982). Cross-cultural application of the Revised-Children's Manifest Anxiety Scale: Normative and reliability data for Nigerian primary school children. *Psychological Reports*, 51(3f), 1135-1138.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 37(1), 51-87.
- 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.
- Prior, M., Sanson, A., Freethy, C., & Geffen, G. (1985). Auditory attentional abilities in hyperactive children. *Journal of Child Psychology and Psychiatry*, 26(2), 289-304.
- Quinn, P. O. (2008). Attention-deficit/hyperactivity disorder and its comorbidities in women and girls: an evolving picture. *Current Psychiatry Reports*, 10(5), 419-423.

- Reynolds, C. R. (1980). Concurrent validity of "What I think and feel:" the Revised Children's Manifest Anxiety Scale. *Journal of Consulting and Clinical Psychology, 48*(6), 774-775.
- Reynolds, C. R. (1981). Long-term stability of scores on the Revised-Children's Manifest Anxiety Scale. *Perceptual and Motor Skills, 53*(3), 702-702.
- Reynolds, C. R. (1982). Convergent and divergent validity of the Revised Children's Manifest Anxiety Scale. *Educational and Psychological Measurement, 42*(4), 1205.
- Reynolds, C. R., & Paget, K. D. (1981). Factor analysis of the Revised Children's Manifest Anxiety Scale for Blacks, Whites, males, and females with a national normative sample. *Journal of Consulting and Clinical Psychology, 49*(3), 352.
- Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale (RCMAS): Manual*: Western Psychological Services.
- Rothen, S., Vandeleur, C. L., Lustenberger, Y., Jeanpretre, N., Ayer, E., Gamma, F., . . . Preisig, M. (2009). Parent-child agreement and prevalence estimates of diagnoses in childhood: direct interview versus family history method. *International Journal of Methods in Psychiatric Research, 18*(2), 96-109.
- Rucklidge, J. J. (2010). Gender differences in attention-deficit/hyperactivity disorder. *The Psychiatric Clinics of North America, 33*(2), 357-373.
- Rucklidge, J. J., & Tannock, R. (2001). Psychiatric, psychosocial, and cognitive functioning of female adolescents with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(5), 530-540.
- Sagvolden, T., Aase, H., Zeiner, P., & Berger, D. (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research, 94*(1), 61-71.
- Sbordone, R. J. (2000). *Ecological validity issues in neuropsychological testing*: Brain Injury Source, 4, 10-12.
- Scholwinski, E., & Reynolds, C. R. (1985). Dimensions of anxiety among high IQ children. *Gifted Child Quarterly, 29*(3), 125-130.
- Sciutto, M. J., & Eisenberg, M. (2007). Evaluating the evidence for and against the overdiagnosis of ADHD. *Journal of Attention Disorders, 11*(2), 106-113.
- Sciutto, M. J., Nolfi, C. J., & Bluhm, C. (2004). Effects of child gender and symptom type on referrals for ADHD by elementary school teachers. *Journal of Emotional and Behavioral Disorders, 12*(4), 247.
- Seidman, L. J. (2006). Neuropsychological functioning in people with ADHD across the lifespan. *Clinical Psychology Review, 26*(4), 466-485.

- Seidman, L. J., Biederman, J., Monuteaux, M. C., Valera, E., Doyle, A. E., & Faraone, S. V. (2005). Impact of gender and age on executive functioning: do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Developmental Neuropsychology*, *27*(1), 79-105.
- Seligman, L. D., Ollendick, T. H., Langley, A. K., & Baldacci, H. B. (2004). The utility of measures of child and adolescent anxiety: a meta-analytic review of the Revised Children's Manifest Anxiety Scale, the State-Trait Anxiety Inventory for Children, and the Child Behavior Checklist. *Journal of Clinical Child and Adolescent Psychology*, *33*(3), 557-565.
- 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.
- Smith, D. G., Xiao, L., & Bechara, A. (2012). Decision making in children and adolescents: impaired Iowa Gambling Task performance in early adolescence. *Developmental Psychology*, *48*(4), 1180-1187.
- Sonuga-Barke, E. J. (2003). The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *Neuroscience and Biobehavioral Reviews*, *27*(7), 593-604.
- Sonuga-Barke, E. J. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biological Psychiatry*, *57*(11), 1231-1238.
- Spielberger, C. D., & Edwards, C. D. (1973). *State-trait Anxiety Inventory for Children: STAIC: How I Feel Questionnaire: Professional Manual*: Mind Garden.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* *18*, 643-662.
- Swanson, J., Baler, R. D., & Volkow, N. D. (2011). Understanding the effects of stimulant medications on cognition in individuals with attention-deficit hyperactivity disorder: a decade of progress. *Neuropsychopharmacology*, *36*(1), 207-226.
- Thompson, P. M., Sowell, E. R., Gogtay, N., Giedd, J. N., Vidal, C. N., Hayashi, K. M., . . . Toga, A. W. (2005). Structural MRI and Brain Development. *International Review of Neurobiology*, *67*, 285-323.
- 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.
- Toplak, M. E., Jain, U., & Tannock, R. (2005). Executive and motivational processes in adolescents with Attention-Deficit-Hyperactivity Disorder (ADHD). *Behavioral and Brain Functions*, 1(1), 8.
- Toplak, M. E., Sorge, G. B., Benoit, A., West, R. F., & Stanovich, K. E. (2010). Decision-making and cognitive abilities: A review of associations between Iowa Gambling Task performance, executive functions, and intelligence. *Clinical Psychology Review*, 30(5), 562-581.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 54(2), 131-143.
- Turgeon, L., & Chartrand, É. (2003). Reliability and Validity of the Revised Children's Manifest Anxiety Scale in a French-Canadian Sample. *Psychological Assessment*, 15(3), 378.
- van den Bos, R., Homberg, J., & de Visser, L. (2013). A critical review of sex differences in decision-making tasks: focus on the Iowa Gambling Task. *Behavioural Brain Research*, 238, 95-108.
- van der Oord, S., Prins, P. J., Oosterlaan, J., & Emmelkamp, P. M. (2008). Efficacy of methylphenidate, psychosocial treatments and their combination in school-aged children with ADHD: a meta-analysis. *Clinical Psychology Review*, 28(5), 783-800.
- van Duijvenvoorde, A. C., Jansen, B. R., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: comparing informed and noninformed situations. *Developmental Psychology*, 48(1), 192-203.
- Wahlstedt, C., Thorell, L. B., & Bohlin, G. (2009). Heterogeneity in ADHD: neuropsychological pathways, comorbidity and symptom domains. *Journal of Abnormal Child Psychology*, 37(4), 551-564.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence. Norwegian version*. Stockholm: The Psychological Corporation.
- Wechsler, D. (2004). *Wechsler Intelligence Scale for Children-Fourth edition. Norwegian version*. Stockholm: The Psychological Corporation.
- 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.

- Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., . . . Lahey, B. B. (2012). Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology, 121*(4), 991-1010.
- Wodka, E. L., Loftis, C., Mostofsky, S. H., Prahme, C., Larson, J. C., Denckla, M. B., & Mahone, E. M. (2008). Prediction of ADHD in boys and girls using the D-KEFS. *Archives of Clinical Neuropsychology, 23*(3), 283-293.
- Øie, M., & Hugdahl, K. (2008). A 10–13 year follow-up of changes in perception and executive attention in patients with early-onset schizophrenia: A dichotic listening study. *Schizophrenia Research, 106*(1), 29-32.
- Øie, M., Rund, B. R., Sundet, K., & Bryhn, G. (1998). Auditory laterality and selective attention: normal performance in patients with early-onset schizophrenia. *Schizophrenia Bulletin, 24*(4), 643-652.
- Øie, M., Sundet, K., & Ueland, T. (2011). Neurocognition and functional outcome in early-onset schizophrenia and attention-deficit/hyperactivity disorder: a 13-year follow-up. *Neuropsychology, 25*(1), 25-35.

## APPENDIX

### Diagnostic criteria for DSM-IV-TR: Attention-Deficit/Hyperactivity Disorder (ADHD)

A. Either (1) or (2):

- 1) six (or more) of the following symptoms of **inattention** have persisted for at least 6 months to a degree that is maladaptive and inconsistent with developmental level:

#### *Inattention*

- a) often fails to give close attention to details or makes careless mistakes in schoolwork, work, or other activities
- b) often has difficulty sustaining attention in tasks or play activities
- c) often does not seem to listen when spoken to directly
- d) often does not follow through on instructions and fails to finish school work, chores, or duties in the workplace (not due to oppositional behavior or failure to understand instructions)
- e) often has difficulty organizing tasks and activities
- f) often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (such as schoolwork or homework)
- g) often loses things necessary for tasks or activities (e.g., toys, school assignments, pencils, books, or tools)
- h) is often easily distracted by extraneous stimuli
- i) is often forgetful in daily activities

- 2) six (or more) of the following symptoms of **hyperactivity-impulsivity** have persisted for at least 6 months to a degree that is maladaptive and inconsistent with developmental level:

#### *Hyperactivity*

- a) often fidgets with hands or feet or squirms in seat
- b) often leaves seat in classroom or in other situations in which remaining seated is expected
- c) often runs about or climbs excessively in situations in which it is inappropriate (in adolescents or adults, may be limited to subjective feelings of restlessness)

- d) often has difficulty playing or engaging in leisure activities quietly
- e) is often "on the go" or often acts as if "driven by a motor"
- f) often talks excessively

Impulsivity

- g) often blurts out answers before questions have been completed
- h) often has difficulty awaiting turn
- i) often interrupts or intrudes on others (e.g., butts into conversations or games)

- B. Some hyperactive-impulsive or inattentive symptoms that caused impairment were present before age 7 years.
- C. Some impairment from the symptoms is present in two or more settings (e.g., at school [or work] and at home).
- D. There must be clear evidence of clinically significant impairment in social, academic, or occupational functioning.
- E. The symptoms do not occur exclusively during the course of a Pervasive Developmental Disorder, Schizophrenia, or other Psychotic Disorder and are not better accounted for by another mental disorder (e.g., Mood Disorder, Anxiety Disorder, Dissociative Disorders, or a Personality Disorder).

*Code based on type:*

**314.01 Attention-Deficit/Hyperactivity Disorder, Combined Type:** if both Criteria A1 and A2 are met for the past 6 months

**314.00 Attention-Deficit/Hyperactivity Disorder, Predominantly Inattentive Type:** if Criterion A1 is met but Criterion A2 is not met for the past 6 months

**314.01 Attention-Deficit/Hyperactivity Disorder, Predominantly Hyperactive-Impulsive Type:** if Criterion A2 is met but Criterion A1 is not met for the past 6 months

















RESEARCH ARTICLE

Open Access

# ADHD in girls and boys – gender differences in co-existing symptoms and executive function measures

Erik Winther Skogli<sup>1\*</sup>, Martin H Teicher<sup>2</sup>, Per Normann Andersen<sup>1</sup>, Kjell Tore Hovik<sup>1</sup> and Merete Øie<sup>1,3</sup>

## Abstract

**Background:** ADHD is diagnosed and treated more often in males than in females. Research on gender differences suggests that girls may be consistently underidentified and underdiagnosed because of differences in the expression of the disorder among boys and girls. One aim of the present study was to assess in a clinical sample of medication naïve boys and girls with ADHD, whether there were significant gender x diagnosis interactions in co-existing symptom severity and executive function (EF) impairment. The second aim was to delineate specific symptom ratings and measures of EF that were most important in distinguishing ADHD from healthy controls (HC) of the same gender.

**Methods:** Thirty-seven females with ADHD, 43 males with ADHD, 18 HC females and 32 HC males between 8 and 17 years were included. Co-existing symptoms were assessed with self-report scales and parent ratings. EF was assessed with parent ratings of executive skills in everyday situations (BRIEF), and neuropsychological tests. The three measurement domains (co-existing symptoms, BRIEF, neuropsychological EF tests) were investigated using analysis of variance (ANOVA) and random forest classification.

**Results:** ANOVAs revealed only one significant diagnosis x gender interaction, with higher rates of self-reported anxiety symptoms in females with ADHD. Random forest classification indicated that co-existing symptom ratings was substantially better in distinguishing subjects with ADHD from HC in females (93% accuracy) than in males (86% accuracy). The most important distinguishing variable was self-reported anxiety in females, and parent ratings of rule breaking in males. Parent ratings of EF skills were better in distinguishing subjects with ADHD from HC in males (96% accuracy) than in females (92% accuracy). Neuropsychological EF tests had only a modest ability to categorize subjects as ADHD or HC in males (73% accuracy) and females (79% accuracy).

**Conclusions:** Our findings emphasize the combination of self-report and parent rating scales for the identification of different comorbid symptom expression in boys and girls already diagnosed with ADHD. Self-report scales may increase awareness of internalizing problems particularly salient in females with ADHD.

**Keywords:** ADHD, Gender, Comorbidity, Executive functions, BRIEF

## Background

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common childhood neuropsychiatric disorders, characterized by problems with inattention, hyperactivity and impulsivity [1]. Worldwide prevalence estimates for childhood ADHD range between 3% and 7% [1] with a male-to-female ratio of 3:1 in population based studies

[2,3] and between 5:1 to 9:1 in clinical samples [1,3,4]. Research on gender differences suggests that girls may be consistently underidentified and underdiagnosed mostly explained by differences in the expression of the disorder among boys and girls [3,5-7]. Females with ADHD are reported to have fewer hyperactive/impulsive symptoms and more inattentive symptoms when compared with males with ADHD [3,8,9]. Further, females with ADHD present more commonly with the inattentive subtype than do boys [10]. Less disruptive behavior in females with ADHD may contribute to referral bias causing underidentification and

\* Correspondence: erik.skogli@sykehuset-innlandet.no

<sup>1</sup>Innlandet Hospital Trust Lillehammer, Division Mental Health Care, Oslo, Norway

Full list of author information is available at the end of the article

lack of treatment for females with ADHD [11]. For example, Sciotto, Nolfi, & Bluhm [12] found that teachers more often refer boys than girls for treatment for ADHD, even when showing equal levels of impairment.

Another major contributing factor to late or missed diagnoses in females appears to be the presence of co-existing symptoms that often cloud the diagnostic picture [5]. As many as 75% of children with ADHD are likely to have at least one other psychiatric disorder [2,13]. Thus, in clinical practice, co-existing psychiatric problems are the rule rather than the exception. Boys with ADHD have been found to have more externalizing disorders than normal developing boys, while females tend to show more internalizing disorders in comparison to normal developing girls [3,5,9]. In addition, adolescent females with ADHD are reported to show more internalizing symptoms than their male counterparts [14]. Often other diagnoses are made long before the diagnosis of ADHD is assessed in female clinical populations because of less overt ADHD symptoms [15]. By comparison, more overt acting out behavior seems to drive referral for ADHD assessment in boys [3].

With regard to executive functions (EF), which are considered a central source of the disability associated with ADHD [16-19], previous research has yielded more similarities than differences between girls and boys with ADHD [11,20]. Despite the centrality of EF deficits in ADHD, a neuropsychological profile distinct to females with ADHD when compared with male counterparts has yet to be identified. In addition, neuropsychological tests have shown to be weakly related to the severity of ADHD symptoms [21], and recent studies suggest that many subjects with ADHD perform normally on EF tests [22,23]. Where traditional neuropsychological EF tests seem to capture "best estimates" in an ideal setting [24], the Behavior Rating Inventory of Executive Function (BRIEF) was designed to assess EF performance in ecologically valid situations [25]. This instrument has shown consistent, but modest correlations with neuropsychological EF tests [26-28]. The BRIEF has proven to be a valuable additional assessment tool for the identification of ADHD in school-aged children [28,29], demonstrating better sensitivity than neuropsychological EF tests [28]. To the best of our knowledge, no studies have investigated potential gender effects on BRIEF in children and adolescents with ADHD. However, Huizinga and colleagues [30] reported elevated levels of executive problems assessed with BRIEF in normal developing boys compared to normal developing girls (age range 5-18 years). These findings are comparable to the data presented in the original version of the BRIEF, showing superior performance in girls compared to boys in general [25].

In sum, the gender gap in clinical populations of subjects with ADHD continues to hamper the correct diagnosis and treatment of females with ADHD. To our knowledge, the current study is the first to include both self-report and

parent ratings of co-existing symptoms in addition to laboratory testing and inventory based scales assessing EF in medication naïve boys and girls with ADHD to examine potential gender sensitive ADHD profiles.

Our first aim was to assess whether there were significant gender x diagnosis interactions in co-existing symptom severity and EF impairment. Using conventional ANOVAs we hypothesised that boys and girls with ADHD would show greater impairment in all three measurement domains (co-existing symptoms, neuropsychological EF tests, BRIEF) relative to HC. However, we expected to find few significant gender x diagnosis interactions in the three measurement domains. Our second aim was to delineate specific symptom ratings and measures of EF that were most important in distinguishing ADHD from HC in the same gender. We used random forest classification with cross-validation, where the identification of subtle differences across diagnosis and gender in moderately sized samples is possible. It was hypothesised that co-existing internalizing symptoms would better distinguish subjects with ADHD from HC in females than in males. Co-existing externalizing symptoms would better distinguish subjects with ADHD from HC in males than in females. Second, neuropsychological test results were expected to distinguish subjects with ADHD from HC equally well in males and females. Finally, parent ratings of EF were hypothesised to better distinguish subjects with ADHD from HC in males than in females.

## Methods

### Procedure and participants

Demographic characteristics are presented in Table 1. Forty three males with ADHD ( $M = 11.2$  years), 37 females with ADHD ( $M = 11.9$  years), 32 healthy control (HC) males ( $M = 11.4$  years) and 18 HC females ( $M = 11.9$  years) between 8 and 17 years participated in the study. The ADHD participants were recruited as consecutive referrals from seven outpatient Child and Adolescent Mental Health Centres in Innlandet Hospital Trust (IHT) for assessment of ADHD. All participants underwent a comprehensive assessment according to common clinical practice. Semi-structured clinical interviews (Kiddie-Schedule for Affective Disorders and Schizophrenia - K-SADS) [31] were conducted separately for children/adolescents and parents to assess psychopathology. The interviewers were experienced clinicians, and were trained to high levels of interrater reliability for the assessment of diagnosis. The diagnostic evaluation with K-SADS was supplemented with information from the ADHD Rating Scale IV (ARS-IV) [32], and the Child Behavior Checklist/6-18 [33], which covers the DSM-IV symptoms for ADHD. Teacher reports describing school functioning, both academic and socially, which is mandatory on referral, were incorporated into the diagnostic evaluation. Diagnoses were considered positive



**Table 1 Demographic characteristics: means and standard deviations within the four groups**

Variable	ADHD		Healthy controls		Group comparison		Bonferroni*
	Boys <sup>1</sup> (n = 43)	Girls <sup>2</sup> (n = 37)	Boys <sup>3</sup> (n = 32)	Girls <sup>4</sup> (n = 18)	F	P	
	Age (months)	139.2 (23.2)	149.4 (25.1)	141.5 (22.6)	148.5 (27.3)	$F(3,126) = 1.50$	
Mother's education (yrs)	12.6 (2.3)	12.9 (1.9)	14.6 (2.5)	14.6 (2.1)	$F(3,126) = 6.80$	< .001	1 < 3,4; 2 < 3,4
FSIQ (WASI) <sup>a)</sup>	94.3 (13.2)	96.4 (15.5)	101.9 (12.7)	107.1 (13.1)	$F(3,126) = 4.60$	= .004	1,2 < 4
Inattention <sup>b)</sup>	16.6 (5.8)	15.0 (5.6)	1.8 (2.1)	1.4 (1.6)	$F(3,125) = 95.50$	< .001	1,2 < 3,4
Hyperactivity/Impulsivity <sup>c)</sup>	10.2 (6.5)	10.1 (6.7)	1.1 (1.3)	0.8 (1.4)	$F(3,125) = 30.86$	< .001	1,2 < 3,4
CGAS <sup>d)</sup>	56.0 (8.4)	59.3 (9.3)			$t(78) = 2.82$	NS	

<sup>a)</sup>Full Scale IQ.

<sup>b)</sup>ADHD rating scale - IV.

<sup>c)</sup>ADHD rating scale - IV.

<sup>d)</sup>Children Global Assessment Scale.

\*1 = ADHD Boys, 2 = ADHD Girls, 3 = HC Boys, 4 = HC Girls.

if, based on a comprehensive evaluation of K-SADS, teacher information and rating scales, DSM-IV [1] criteria were met.

Based on diagnostic evaluation with K-SADS, co-existing diagnoses within the group of males with ADHD included depression (4.7%), anxiety (4.7%), conduct disorder (4.7%), and oppositional defiant disorder (11.6%). Co-existing diagnoses within the females with ADHD included anxiety (8.1%), and oppositional defiant disorder (10.8%). Despite a low prevalence of co-existing diagnoses, parent and self-report scales indicated elevated levels of externalizing and internalizing symptoms in both males and females with ADHD when compared with normal developing counterparts. Exclusion criteria for all participants included prematurity (< 36 weeks), IQ below 70, a history of stimulant treatment or any disease affecting the central nervous system. None of the participants used any type of psychopharmacological medication. One boy with ADHD was excluded due to estimated IQ below 70. None were excluded due to history of stimulant treatment or any disease affecting the central nervous system.

All participants in the HC groups were screened for mental disorders with the K-SADS in separate interviews for children/adolescents and parents. The HC were recruited from local schools and were given a small compensation for participating. The HC could not have been treated for a mental disorder, have a psychiatric diagnosis, have had a head injury (with loss of consciousness) or known dyslexia. The four groups (ADHD/females, ADHD/males, HC/females, HC/males) did not differ significantly with regard to age and gender distribution. The Wechsler Abbreviated Scale of Intelligence (WASI) [34] was administered to estimate IQ in all participants. The groups differed significantly with regard to IQ,  $F(3,126) = 4.60$ ,  $p = .004$ ,  $\eta^2 = .099$ , and Bonferroni post-hoc analysis showed that both ADHD gender groups scored below the females in the HC group. On average, mothers of children in the HC

group had 1.7 years more education than mothers of children with ADHD,  $F(3,126) = 6.80$ ,  $p < 0.001$ . All parents/caregivers and participants above 12 years gave written informed consent in accordance with the Research Ethics Committee in Eastern Norway. All children under the age of 12 years provided oral consent to participate. The study was approved by the Regional Committee for Medical Research Ethics in Eastern Norway (REK-Øst), and by the Privacy protection ombudsman for research at Innlandet Hospital Trust. It was conducted in accordance with the Helsinki Declaration of the World Medical Association Assembly.

## Measures

### Measures of symptomatology

The Child Behavior Checklist/6-18 (CBCL) [33] is a widely used scale containing 7 competence items and 113 specific problem items, each of which is rated on a 0–2 metric. The 120 items assess adaptive behavior as well as eight narrow band factors (Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior and Aggressive Behavior) and two broadband factors (Externalizing and Internalizing symptoms) of co-existing symptoms. The 2001 revision also includes seven DSM-oriented scales consistent with DSM diagnostic categories (Affective Problems, Anxiety Problems, Somatic Problems, ADHD, Oppositional Defiant Problems and Conduct Problems). On the parent-report CBCL, we used seven of the narrow band factors (excluding Attention Problems) and five of the DSM-oriented scales (excluding ADHD) to assess co-existing symptoms. Elevated T-scores indicate a higher degree of co-existing internalizing and externalizing symptoms. Cross-cultural studies have demonstrated good discriminant validity with mean factor loadings across societies at .62 [35]. Acceptable reliability and validity

of the Norwegian version of the CBCL are reported by Novik [36,37].

The Revised Children's Manifest Anxiety Scale, second edition (RCMAS-2) [38] is a 49-item self-report instrument designed to measure anxiety symptoms in children 6 to 19 years of age. Children respond either "Yes" or "No" to all 49-items. The instrument reveals three anxiety factors: Physiological Anxiety, Worry and Social Anxiety. The three anxiety factors are summed yielding a Total Anxiety score. Elevated raw-scores indicate a higher degree of anxiety symptoms. The RCMAS Total Anxiety Scale has been found to have satisfactory psychometric properties with high test-retest reliability [39,40] and consistent construct validity [41-44]. Satisfactory psychometric properties have been replicated among other cultures as well [39,45-47].

The State-Trait Anxiety Inventory for Children (STAIC) [48] includes two 20-item self-report scales that measure both enduring tendencies (Trait) and situational variations (State) in levels of perceived anxiety. Children respond on a three-point scale indicating varying degree of worry, feelings of tension, and/or nervousness. Elevated raw-scores indicate a higher degree of situational and temporal anxiety. In a quantitative review by Seligman and colleagues [49], the authors argue that the STAIC possess satisfactory psychometric properties.

The Short Mood and Feelings Questionnaire (SMFQ) [50] is a 13-item self-report instrument designed to measure depressive symptoms in children 8 to 18 years of age. The SMFQ is derived from the original 30-item Mood and Feelings Questionnaire (MFQ) [51] where children respond on a three-point scale ("not true", "sometimes true" and "true"). A net score was generated based on the 13 items with elevated raw-scores indicating a higher degree of depression symptoms. The SMFQ have demonstrated high internal consistency (Cronbach's  $\alpha = .90$ ) [52], and test-retest stability in children for a two-week period yielded an intra class correlation of .66 [51]. Angold and colleagues [50] found SMFQ to correlate strongly with Children's Depression Inventory (CDI) [53] and Diagnostic Interview Schedule for Children (DISC-C) depression scores [51] ( $r = .67$  and  $.51$ , respectively).

### **Neuropsychological EF tests**

#### **The letter-number sequencing test**

The Letter-Number Sequencing Test (LN) [54] was used as a measure of working memory. The test consists of ten items. Each item contains three trials with the same number of digits and letters. The test administrator reads aloud each trial and asks the child to recall the numbers in ascending order and the letters in alphabetical order. In the present study, total correct recalled trials were examined. Lower scaled scores indicated difficulties with the task.

#### **The colour - word interference test, condition 3**

The Colour - Word Interference Test, Condition 3 (CW 3) [55,56] was used as a measure of inhibition. The examinee needs to inhibit an overlearned verbal response when naming the dissonant ink colours in which the words are printed. For the present study, completion time in seconds was examined. Lower scaled scores indicated difficulties with the task.

#### **The colour - word interference test, condition 4**

The Colour - Word Interference Test, Condition 4 (CW 4) [56] was used as a measure of cognitive flexibility. The examinee is asked to switch back and forth between naming the dissonant ink colours and reading the words. For the present study, completion time in seconds was examined. Lower scaled scores indicated difficulties with the task.

#### **The trail making test, condition 4**

The Trail Making Test, condition 4 (TMT 4) [56] was used as a measure of cognitive flexibility. The examinee is asked to draw a line interchangeably between numbers and letters in the right order. For the present study, time to complete task was examined. Lower scaled scores indicated difficulties with the task.

#### **The design fluency test, condition 3**

The Design Fluency Test, condition 3 (DF) [56] was used as a measure of cognitive flexibility. The examinee is asked to draw as many different designs as possible using four straight lines connecting five filled and empty dots interchangeably. The examinee is given 60 seconds for the task. For the present study, total correct responses were examined. Lower scaled scores indicated difficulties with the task.

#### **The tower test**

The Tower Test [56] was used as a measure of planning. In this task the examinee is asked to construct several target towers by moving five disks, varying in size, across three pegs in the fewest number of moves possible. While doing this, the examinee is allowed to move only one disk at a time, and not to place a larger disk over a smaller disk. In the present study total achievement score was examined. Lower scaled scores indicated difficulties with the task.

#### **The letter fluency test**

The Letter Fluency Test (LF) [56] was used as a measure of verbal fluency. This task includes three 60-seconds trials, where participants were asked to generate words fluently in an effortful, phonemic format with the letters F, A, and S. For the present study, total correct responses were

examined. Lower scaled scores indicated difficulties with the task.

#### **Inventory based information of EF**

The BRIEF for children and adolescents aged 5 to 18 includes a parent form and a teacher form [25]. In the current study, the Norwegian parent rating version was used. The BRIEF is composed of eight clinical scales (Inhibition, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials and Monitor). Fallmyr & Egeland [57] reported high internal consistency (Chronbachs  $\alpha = .76-.92$ ) on the Norwegian parent rating version of the BRIEF. These values are at the same level as Chronbachs  $\alpha$  reported in the BRIEF manual (.80-.98) [25]. Elevated BRIEF T-scores indicate a higher degree of impairment.

#### **Data analyses**

Data analyses were conducted using the statistical package SPSS for Windows, version 15.0 (SPSS, Inc., Chicago, IL). Demographic characteristics were investigated using the Chi-square test for independence (nominal variables) and analysis of variance (ANOVA) (continuous variables) followed up by Bonferroni post-hoc tests for group comparisons when adequate. ANOVAs were carried out to investigate gender x diagnosis interactions in the three measurement domains (co-existing symptoms, neuro-psychological EF tests, BRIEF).

#### **Random forest classification**

In addition to tests of significance we also used an algorithmic modelling/data mining technique to explore gender differences in co-existing symptoms and EF ratings and measures. Classical statistical techniques are designed to test and reject the hypothesis that observed differences between groups have occurred by chance. Algorithmic modelling techniques have been developed to address a somewhat different question. Briefly, these techniques can identify from a sample of potential predictor variables the most important subset for categorizing subjects or predicting outcomes [58]. Hence, we used this approach to delineate within each gender the subset of symptom ratings, EF measures and EF ratings that appear to be most important in discriminating children with ADHD from HC. Specifically, we used random forest classification and cross-validation (R packages *randomForest* 4.5-34 and *caret* 5.02-011) [59] to identify and rank order different symptom ratings and EF measures for their degree of importance in differentiating ADHD from HC within each gender. Although importance and statistical significance often go hand-in-hand, the two are not necessarily the same. The approach has many advantages. In particular, it can provide meaningful results with smaller sample sizes than stochastic models. Further they are less susceptible

to overfitting and multicollinearity, provide more accurate predictions, and do not make the unlikely assumption that the multivariate data being analyzed are multivariate normal.

Briefly, this is a form of "ensemble learning" in which a large number of unpruned decision trees are generated and their results aggregated [60]. The random part comes in as each tree is constructed using a different bootstrap sample of the data, and each node is split using the best among a subset of predictors randomly chosen at that node. As Liaw and Wiener indicate [61] this strategy performs very well compared to many other classifiers, including discriminant analysis, logistic regression, support vector machines and neural networks [60]. It is primarily used in data mining and in genomic analysis, such as microarray studies.

Each decision tree was generated using results from 75% of the participants and then tested on the remaining 25% (validation set). This process was performed 5000 times on different random splits of the data to provide a cross-validated estimate of the predictive discriminant ability of the measures (accuracy, kappa) that would likely generalize to new cases [62]. The importance of each variable in the cluster was assessed by calculating the decrease in predictive accuracy following the sequential permutation (effective randomization and elimination) of each variable in the cluster on the validation set. The most important variables were the ones whose effective elimination from the forest produced the greatest degradation in accuracy.

#### **Results**

The first statistical approach tested the hypothesis that there were significant gender x diagnosis interactions across the array of dependent variables (DVs), viewing each of the DVs in isolation. In the second approach we evaluated the ability of ratings or measures in the: (1) symptom (CBCL, RCMAS-2, STAIC, SMFQ), (2) EF test performance, and (3) EF rating (BRIEF) clusters to predict whether participants met criteria for ADHD, and if the most important predictor variables in each cluster were the same for males and females.

As seen in Table 2 there were marked group differences between participants with ADHD and HC in symptom ratings. In general, there was a roughly parallel increase in symptom ratings with diagnosis across gender, and the only significant diagnosis x gender interaction was observed in ratings of physiological anxiety on the RCMAS-2. There were significant main effects of diagnosis on several of the EF measures including: working memory (LN), inhibition (CW3), and cognitive flexibility (CW4, TMT4, DF). However, there were no significant gender x diagnosis interactions on these measures (Table 3). Interestingly, there were also robust group differences between parent ratings of

**Table 2 Group differences and interaction effects on symptom ratings (ANOVA): means and standard deviations**

	Group				Main and interaction effects						
	ADHD		Healthy controls		Group		Gender		Group X Gender		
	Boys (n = 43)	Girls (n = 37)	Boys (n = 32)	Girls (n = 18)	F	p	F	p	F	p	
<b>CBCL<sup>(a)</sup></b>											
Anxious/depressed	58.3 (8.2)	61.1 (10.6)	51.5 (4.0)	51.4 (2.7)	31.11	<.001	1.52	NS	0.98	NS	
Withdrawn/depressed	59.4 (8.2)	58.7 (7.2)	51.5 (2.5)	51.2 (2.6)	45.29	<.001	0.26	NS	0.04	NS	
Somatic complaints	59.6 (8.9)	59.7 (8.6)	53.0 (4.7)	51.4 (2.5)	29.25	<.001	0.30	NS	0.35	NS	
Social problems	60.4 (9.2)	60.0 (7.4)	50.5 (1.5)	50.3 (0.5)	65.08	<.001	0.02	NS	0.00	NS	
Thought problems	58.1 (8.8)	57.2 (7.7)	51.1 (2.1)	50.6 (1.5)	31.40	<.001	0.36	NS	0.02	NS	
Rule-breaking	60.3 (8.6)	58.6 (8.0)	50.9 (2.6)	50.8 (2.5)	52.06	<.001	1.59	NS	0.52	NS	
Aggressive behavior	61.8 (11.7)	61.5 (9.6)	50.9 (1.8)	50.7 (2.1)	47.81	<.001	0.03	NS	0.00	NS	
Affective problems	61.8 (8.7)	63.2 (8.8)	51.4 (3.6)	50.9 (1.7)	72.89	<.001	0.27	NS	0.47	NS	
Anxiety problems	57.0 (7.8)	60.7 (8.4)	51.2 (3.4)	51.2 (3.3)	35.13	<.001	4.08	.045	2.21	NS	
Somatic problems	60.0 (8.6)	60.1 (10.3)	53.4 (5.0)	51.6 (2.9)	25.96	<.001	0.17	NS	0.43	NS	
Oppositional problems	59.7 (8.4)	59.8 (8.3)	51.3 (2.2)	51.5 (2.7)	45.75	<.001	0.00	NS	0.00	NS	
Conduct problems	62.6 (9.8)	60.1 (8.3)	51.2 (2.6)	50.9 (2.4)	61.60	<.001	2.18	NS	0.80	NS	
<b>SMFQ<sup>(b)</sup></b>	6.3 (4.9)	7.7 (4.9)	2.4 (2.8)	1.7 (1.6)	38.83	<.001	0.70	NS	1.91	NS	
<b>RCMAS-2<sup>(c)</sup></b>											
Physiological anxiety	5.1 (2.2)	6.5 (2.3)	2.6 (2.2)	1.4 (1.6)	82.35	<.001	1.61	NS	10.46	.002	
Worry	4.2 (4.0)	6.3 (3.8)	2.1 (1.8)	2.0 (1.9)	25.06	<.001	4.83	.030	3.40	NS	
Social anxiety	3.6 (3.2)	5.4 (2.7)	1.3 (1.6)	1.4 (1.7)	39.42	<.001	5.94	.016	2.95	NS	
<b>STAIC<sup>(d)</sup></b>											
State	29.2 (4.1)	30.1 (5.0)	27.3 (4.0)	27.4 (3.4)	8.21	.005	0.23	NS	0.16	NS	
Trait	30.5 (8.5)	34.4 (7.9)	27.7 (5.2)	26.5 (5.9)	13.42	<.001	2.39	NS	3.67	NS	

<sup>(a)</sup>The Child Behavior Checklist/6-18 (df = 1,124). <sup>(b)</sup>The Short Mood and Feelings Questionnaire (df = 1,123). <sup>(c)</sup>The Revised Children's Manifest Anxiety Scale, second edition (df = 1,124). <sup>(d)</sup>The State-Trait Anxiety Inventory for Children: State (df = 1,121), Trait (df = 1,122).  
 Note. Higher scores denote greater pathology.

**Table 3 Group differences and interaction effects on executive function measures (ANOVA): means and standard deviations**

	Group				Main and interaction effects					
	ADHD		Healthy controls		Group		Gender		Group X Gender	
	Boys (n = 43)	Girls (n = 37)	Boys (n = 32)	Girls (n = 18)	F	p	F	p	F	p
LN <sup>(a)</sup>	8.2 (2.9)	8.6 (2.6)	10.6 (1.5)	10.9 (1.7)	29.39	<.001	0.85	NS	0.01	NS
TMT4 <sup>(b)</sup>	7.1 (3.2)	7.8 (3.3)	9.3 (2.9)	10.2 (2.5)	17.14	<.001	2.24	NS	0.00	NS
CW3 <sup>(c)</sup>	8.4 (3.3)	8.3 (3.4)	10.0 (2.4)	11.5 (1.7)	17.40	<.001	1.11	NS	2.17	NS
CW4 <sup>(d)</sup>	7.9 (2.9)	7.1 (3.5)	10.0 (2.6)	10.9 (1.6)	28.46	<.001	0.07	NS	2.76	NS
Tower <sup>(e)</sup>	9.9 (2.2)	10.6 (2.3)	11.0 (2.1)	10.3 (1.7)	1.93	NS	0.13	NS	3.26	NS
DF <sup>(f)</sup>	8.2 (2.5)	9.8 (3.8)	11.8 (2.9)	12.3 (2.4)	31.51	<.001	4.82	.030	0.83	NS
LF <sup>(g)</sup>	8.2 (0.5)	9.1 (0.3)	11.3 (0.5)	11.8 (0.9)	26.21	<.001	1.53	NS	0.06	NS

<sup>(a)</sup>The Letter-Number Sequencing Test (df = 1,123). <sup>(b)</sup>The Trail Making Test, condition 4 (df = 1,123). <sup>(c)</sup>The Colour-Word Interference Test, condition 3 (df = 1,125).  
<sup>(d)</sup>The Colour-Word Interference Test, condition 4 (df = 1,125). <sup>(e)</sup>The Tower Test (df = 1,124). <sup>(f)</sup>The Design Fluency Test, condition 3 (df = 1,116). <sup>(g)</sup>The Letter Fluency Test (df = 1,123).

children with ADHD and of HC on the BRIEF (Table 4). However, on none of the BRIEF ratings were there significant diagnosis by gender interactions.

Random forest classification investigating the relationship between ADHD status and non-ADHD symptom cluster had a predictive (i.e., cross-validated) diagnostic accuracy of  $0.860 \pm 0.058$  (mean  $\pm$  SD) in males and  $0.932 \pm 0.055$  in females. Kappa coefficients were  $0.715 \pm 0.115$  and  $0.844 \pm 0.133$  for males and females, respectively. This indicates that random forest classification using non-ADHD symptom ratings was substantially better at discriminating female subjects with ADHD from female HC than they were in discriminating males with ADHD from male HC (accuracy:  $z = 5.8, p < 10^{-8}$ ; kappa:  $z = 4.7, p < 10^{-5}$ ; two-sample Z-test). As seen in Figure 1, the rank order pattern of variable importance for discriminating ADHD from HC was similar in males and females ( $r_s = 0.676, p < 0.003$ ). CBCL ratings of social problems and affective problems were key distinguishing variables in both genders. However, rule breaking was the most important distinguishing variable in males, while physiological anxiety symptoms on the RCMAS-2 was the most important distinguishing variable in females.

Random forest classification investigating the relationship between ADHD status and EF tests assessing: working memory, inhibition, cognitive flexibility, planning, and verbal fluency had only a modest ability to distinguish participants with ADHD from HC (males: accuracy =  $0.734 \pm 0.078$ , kappa =  $0.466 \pm 0.152$ ; females: accuracy =  $0.785 \pm 0.078$ , kappa =  $0.507 \pm 0.175$ ). There was a significant gender difference in the predictive ability of random forests based on measures of accuracy but not kappa (accuracy:  $z = 3.22, p < 0.001$ ; kappa:  $z = 1.22, p < 0.12$ ). As seen in Figure 2, the rank order patterning of variable importance

on these measures were not significantly correlated between genders ( $r_s = 0.143, p > 0.7$ ). The most important distinguishing variables in males were performance on tests assessing cognitive flexibility (DF) and verbal fluency (LF). The most important distinguishing variables in females were performance on tests assessing working memory (LN), inhibition (CW 3), and cognitive flexibility (CW 4).

Random forest classification investigating the relationship between ADHD status and BRIEF showed that parent ratings of executive skills were able to discriminate males with ADHD from male HC with high accuracy ( $0.960 \pm 0.036$ , kappa  $0.916 \pm 0.076$ ). Random forest classification using BRIEF items was not quite as accurate in discriminating females with ADHD from female HC (accuracy =  $0.923 \pm 0.051, z = 4.29, p < 10^{-5}$  versus males; kappa =  $0.818 \pm 0.123, z = 4.85, p < 10^{-6}$ ). As seen in Figure 3, BRIEF working memory was the most important distinguishing variable in both genders. However, the rank ordering of importance of the eight BRIEF variables between genders correlated to only a marginal degree ( $r_s = 0.667, p = 0.07$ ).

## Discussion

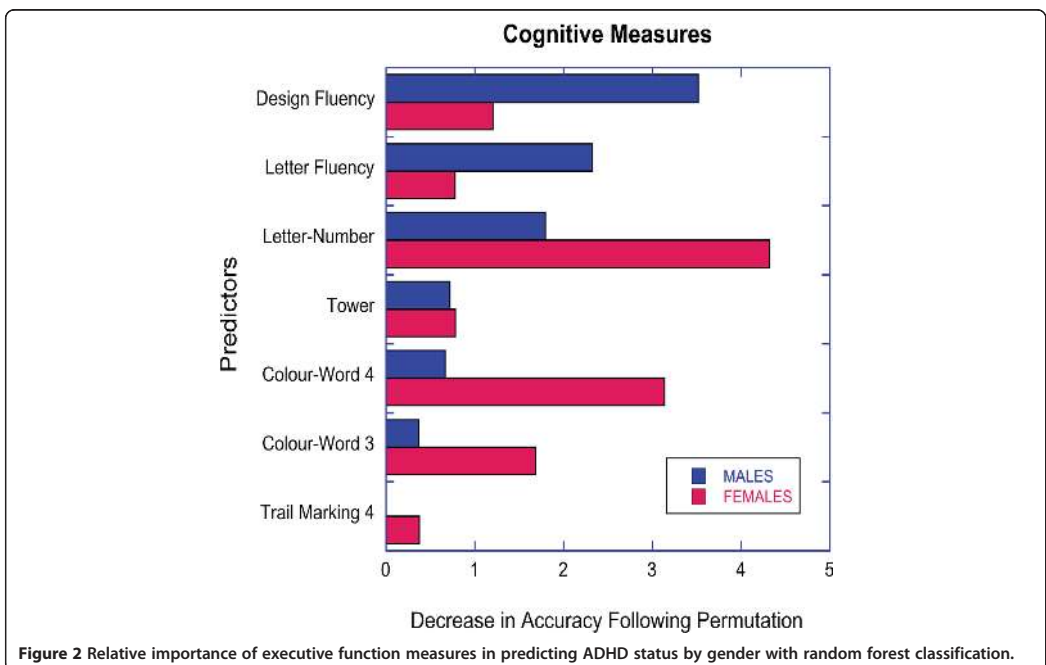
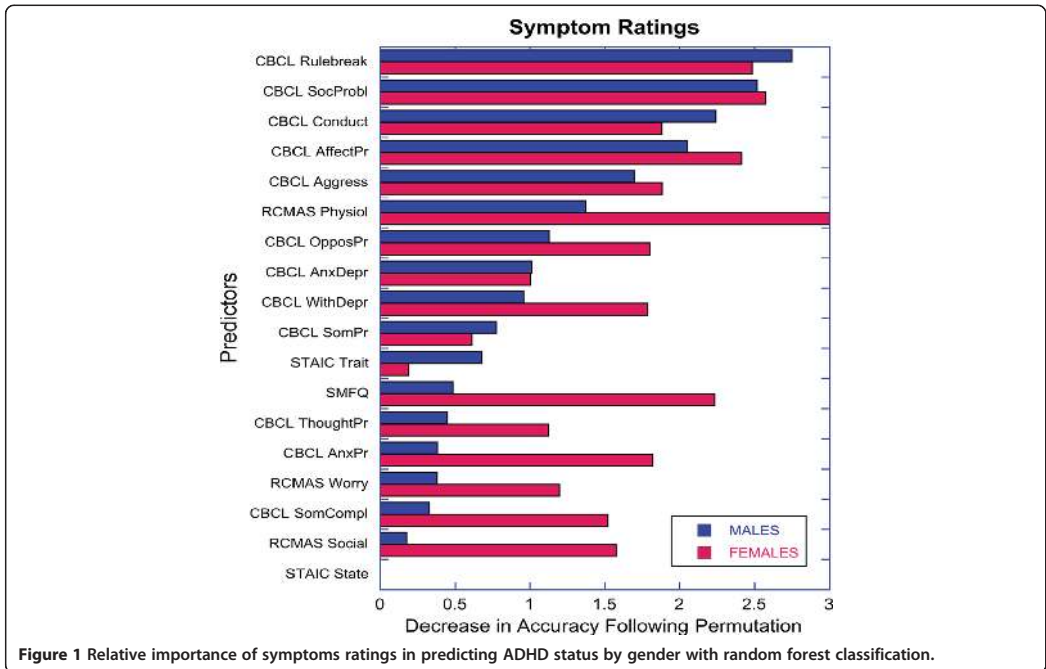
As expected, results on non-ADHD symptom ratings, EF ratings and EF measures differed substantially between ADHD subjects and HC. Boys and girls with ADHD showed in general greater impairment in all three measurement domains relative to HC. There was little evidence for diagnosis  $\times$  gender differences in mean ratings or measures, with more self-reported physiological anxiety in females with ADHD relative to the other groups, as the only significant diagnosis  $\times$  gender finding. Elevated levels of co-existing internalizing symptoms in females with ADHD relative male counterparts is in accordance with

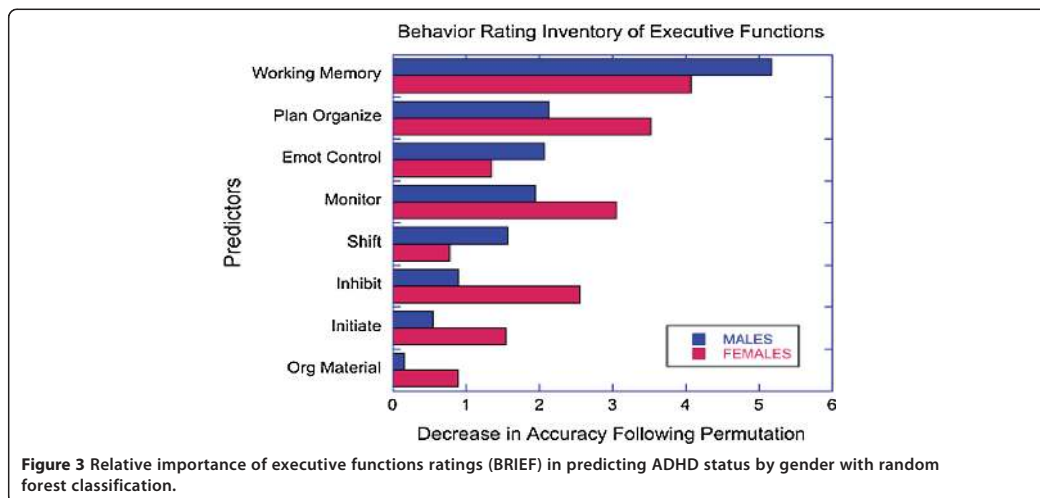
**Table 4 Group differences and interaction effects on executive function ratings (ANOVA): means and standard deviations**

	Group				Main and interaction effects					
	ADHD		Healthy controls		Group		Gender		Group X Gender	
	Boys (n = 43)	Girls (n = 37)	Boys (n = 32)	Girls (n = 18)	F*	p	F*	p	F*	p
Inhibit	59.7 (15.9)	62.5 (12.9)	42.3 (3.5)	42.6 (3.3)	75.17	<.001	0.60	NS	0.31	NS
Shift	57.5 (13.7)	57.3 (11.3)	40.9 (5.8)	40.6 (3.1)	77.65	<.001	0.09	NS	0.00	NS
Emotional control	59.0 (14.1)	61.3 (13.7)	40.7 (3.7)	41.4 (5.3)	88.14	<.001	0.32	NS	0.12	NS
Initiate	59.2 (11.5)	61.8 (12.1)	41.0 (7.2)	40.1 (5.8)	112.18	<.001	0.70	NS	0.96	NS
Working memory	69.3 (9.9)	70.5 (10.1)	41.6 (4.9)	42.8 (4.1)	333.64	<.001	0.56	NS	0.00	NS
Plan/Organize	64.6 (9.8)	67.0 (10.3)	41.4 (5.3)	41.3 (4.1)	250.37	<.001	0.66	NS	0.60	NS
Organization of materials	56.7 (11.4)	58.3 (10.8)	41.3 (7.8)	42.1 (7.4)	75.41	<.001	0.63	NS	0.06	NS
Monitor	60.3 (11.9)	63.8 (12.3)	37.9 (6.1)	39.7 (4.5)	160.68	<.001	1.91	NS	0.17	NS

\**df* = 1,125.

Note: Elevated BRIEF T-scores indicate a higher degree of impairment, with T-scores of 65 and above considered to represent clinically significant areas of concern.





the study by Rucklidge & Tannock [14]. However, few diagnosis x gender differences in general is consistent with previous reports from population-based studies indicating that the disorder is expressed similarly in boys and girls [8,63]. On the other hand, random forest classification with cross-validation provided evidence for meaningful gender differences when investigating the relationship between ADHD status and the three measurement clusters, and in the relative importance of specific items.

First, random forest regression with cross-validation of the symptom cluster indicated that non-ADHD symptoms appeared to be better at categorizing participants as HC or ADHD in females than in males. Our results corroborate previous findings by Rucklidge & Tannock [14] reporting that parent and teachers reported more difficulties with oppositional behaviors, conduct problems, social difficulties, anxiety, and depression in females with ADHD compared to male counterparts. Together these findings provide evidence that co-existing psychological impairment may be even more reliably evident in females with ADHD compared to HC females than in males with ADHD relative to HC males.

Consistent with our hypothesis random forest regression with cross-validation indicated that the most important non-ADHD symptom for categorizing females as ADHD or HC was increased self-reported physiological anxiety (internalizing symptoms), whereas the most important symptom in males was parent rated rule breaking (externalizing symptoms). Elevated ratings of somatic complaints in girls with ADHD have previously been reported in population based studies [64], and supports a hypothesis of somatic complaints as markers for anxiety proneness in females with ADHD [65,66]. Elevated levels of externalizing

symptoms in males with ADHD have been documented in several research reviews [3,5,9].

Third, random forest regression with cross-validation showed that neuropsychological measures of EF had only a modest ability to categorize participants as ADHD or HC. Differences in accuracy when categorizing participants as HC or ADHD in females versus males were slight (79% versus 73% accuracy), and were consistent with our hypothesis of no major difference in discriminatory power between genders. These results corroborate previous findings reporting moderate validity of EF tests for classifying children with ADHD [18,23]. Interestingly, there were differences between genders in the specific EF measures that appeared to be the most important distinguishing variables. Cognitive flexibility and verbal fluency were the most important distinguishing variables in males, whereas working memory and inhibition were the most important distinguishing variables in females. A few previous studies have reported sex differences in EF in children with ADHD [67,68] though most studies report similar EF profiles [11,20]. In sum, EF tests show limited sensitivity and specificity for classifying boys and girls with ADHD.

Fourth, random forest regression with cross-validation indicated that parental ratings of EF were relatively robust distinguishing variables of ADHD status in this sample. In line with our hypothesis, BRIEF scales was significantly better at discriminating males with ADHD from HC males than they were in discriminating females with ADHD from female HC. While there were differences in the relative importance of the different EF ratings, working memory appeared to be the most important distinguishing variable in both boys and girls. The sensitivity of working memory ratings in distinguishing boys and girls with ADHD from

HC has been previously documented by Isquith & Gioia [69] and McCandless & O'Laughlin [29]. Our results indicating better discriminatory power in males than in females with the BRIEF, suggests that ADHD in males may be more reliably associated with impairments in EF than in females. This higher risk of behavioral problems in males with ADHD symptoms may be one reason for the referral of more boys than girls for clinical evaluation of ADHD, as previously suggested by Gaub and Carlson [3].

In our study, the participants' symptoms were assessed with both self-report and parent rating scales. Previous research has shown that clinical samples of children often report more symptoms about themselves than parents do with regard to anxiety and depression [70-74]. In contrast, parents are found to report more conduct disorders [72] or behavioral symptoms than their children [75-77]. Our results with self-reported internalizing symptoms (physiological anxiety) in females with ADHD and parent rated functional impairment in males with ADHD reflect previous reports regarding a self-report/parent rating discrepancy. As internalizing symptoms are less readily observed, parental reports of anxiety and depression are less sensitive than parental reports of externalizing behaviors [77]. It has been speculated that high levels of anxiety and depression in females with ADHD indicate that females are more negatively affected than their male counterparts [78]. The inclusion of self-report scales in clinical practice may thus increase awareness of internalizing problems particularly salient in females with ADHD, and intervention should target co-existing anxiety and depression when indicated. Where parent ratings seem to be informative regarding behavioral problems in boys with ADHD, self-report scales may be more informative regarding internalizing problems in girls with this disorder. Thus, clinical intervention should be sensitive to different expression in co-existing symptoms for boys and girls with ADHD, in addition to conventional treatment of ADHD symptoms.

Strengths of the present study are inclusion of subjects not medicated with psychopharmaca and no history of stimulant treatment. Additional strengths are the use of both self-report and parent ratings when assessing co-existing symptoms, and laboratory tests and inventory-based scales assessing EF. Further, random forest classification [60] is a relatively novel method of determining variable importance, with the advantages of very high classification accuracy and no restrictions regarding the distribution and scaling properties of the data [79]. These properties make random forest regressions well-suited for the classification of large sets of data. It is interesting in this context that predictive modelling with random forests identified gender differences in the accuracy of cluster based categorizations and in the importance of specific measures, whereas conventional statistical analyses showed

only one instance of a significant diagnosis by gender interactions. This is basically a consequence of limited power of the ANOVA test to detect small or subtle interactions. Significant interactive effects are typically detected when the opposite response pattern is seen across gender, or when a large effect is present in one gender and a small effect is present in the other. In order to detect a subtle interactive effect ( $f = 0.1$ ) with power of 0.8 would have required a much larger sample size ( $n = 787$ ).

We used random forests with cross-validation to simultaneously assess clusters of variables to delineate predictive accuracy and to identify the most important distinguishing variables. This approach has been shown to be effective in identifying the most important predictors in large sets of variables with fewer participants than can even be considered using conventional statistical techniques [60]. For instance, this approach can be used in microarray studies to accurately identify the most important subset of polymorphisms even when the number of variables greatly exceeds the number of participants. In short, this approach makes it possible to identify subtle differences across diagnosis and gender in symptom measures and ratings in moderately sized samples.

The present findings need to be interpreted in the context of some methodological limitations. Participants were recruited from a sample of referred subjects, and consequently are not necessarily representative of children and adolescents with ADHD in the general population. However, we believe the participants are fairly representative of males and females referred for the evaluation of ADHD related symptoms. Previous findings have reported that females with ADHD present more commonly with the inattentive subtype than do boys [10]. These subtype effects may potentially have an impact on gender dependent symptom profiles observed in our study. However, despite an overload of males in the ADHD-C group (23 males, 13 females), subtype distribution did not differ significantly by gender in our sample, and the level of inattention and hyperactivity/impulsivity symptoms (ARS-IV) [32] was equal between males and females with ADHD. Albeit previous studies have reported gender differences in hyperactive/impulsive and inattentive symptoms [3,8,9], findings by Lambek and colleagues [80] indicated that these gender dependent subtype effects may be more evident in non-referred than in referred samples of boys and girls with ADHD.

## Conclusions

Overall, females with ADHD could be more accurately distinguished from HC by the presence of co-existing symptoms, particularly physiological components of anxiety. On the other hand, parental reports of EF impairments were substantially better at distinguishing subjects with ADHD from HC in males. Given the almost universal



phenomenon of “co-morbidity” in ADHD, our findings emphasize the combined value of self-report and parent rating scales for the identification of comorbid symptoms in boys and girls already diagnosed with ADHD.

#### Competing interests

The authors declare no conflict of interest with respect to authorship or publication of this article.

Parts of this paper were presented as a poster at the Eunethydis 2nd International ADHD Conference in Barcelona 23–25 May 2012.

#### Authors' contributions

EWS managed the literature searches, undertook the preliminary statistical analyses, interpreted the data, and wrote the first draft of the current manuscript. MHT undertook the statistical analyses, helped interpret the data, and revised the current manuscript critically. PNA helped collect the data, and revised the current manuscript critically. KTH helped collect the data, and revised the manuscript critically. MØ wrote the protocol, managed the literature searches, helped interpret the data and revised the current manuscript critically. All consented to their names on the final manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

The project has received financial support from Innlandet Hospital Trust (grant number 150186) and the Regional Resource Center for Autism, ADHD, Tourette syndrome and Narcolepsy, Oslo University Hospital (grant number 150182). Dr. Teicher's efforts were supported by US NIH award RC1 MH089743.

#### Author details

<sup>1</sup>Innlandet Hospital Trust Lillehammer, Division Mental Health Care, Oslo, Norway. <sup>2</sup>Harvard Medical School, Boston, USA. <sup>3</sup>Department of Psychology, University of Oslo, Innlandet Hospital Trust Lillehammer, Division Mental Health Care, Oslo, Norway.

Received: 4 November 2012 Accepted: 5 November 2013

Published: 9 November 2013

#### References

1. American Psychiatric Association: *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. 4th edition. Washington, DC: American Psychiatric Association; 2000.
2. Barkley RA: *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment*. The Guilford Press; 2006.
3. Gaub M, Carlson CL: Gender differences in ADHD: a meta-analysis and critical review. *J Am Acad Child Adolesc Psychiatry* 1997, **36**(8):1036–1045.
4. Sandberg S: *Hyperactivity and attention disorders of childhood*. 2nd edition. Cambridge, England: Cambridge University Press; 2002.
5. Quinn PO: Attention-deficit/hyperactivity disorder and its comorbidities in women and girls: an evolving picture. *Curr Psychiatry Rep* 2008, **10**(5):419–423.
6. Biederman J, Mick E, Faraone SV, Braaten E, Doyle A, Spencer T, Wilens TE, Frazier E, Johnson MA: Influence of gender on attention deficit hyperactivity disorder in children referred to a psychiatric clinic. *Am J Psychiatry* 2002, **159**(1):36–42.
7. Scituito MJ, Eisenberg M: Evaluating the evidence for and against the overdiagnosis of ADHD. *J Atten Disord* 2007, **11**(2):106–113.
8. Biederman J, Kwon A, Aleardi M, Chouinard VA, Marino T, Cole H, Mick E, Faraone SV: Absence of gender effects on attention deficit hyperactivity disorder: findings in nonreferred subjects. *Am J Psychiatry* 2005, **162**(6):1083–1089.
9. Gershon J: A meta-analytic review of gender differences in ADHD. *J Atten Disord* 2002, **5**(3):143–154.
10. Hinshaw SP, Owens EB, Sami N, Fargeon S: Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *J Consult Clin Psychol* 2006, **74**(3):489–499.
11. Rucklidge JJ: Gender differences in attention-deficit/hyperactivity disorder. *Psychiatr Clin North Am* 2010, **33**(2):357–373.
12. Scituito MJ, Nolfi CJ, Bluhm C: Effects of child gender and symptom type on referrals for ADHD by elementary school teachers. *J Emotion Behav Disord* 2004, **12**(4):247–253.
13. Barkley RA, Murphy KR, Fischer M: *ADHD in adults: What the science says*. The Guilford Press; 2010.
14. Rucklidge JJ, Tannock R: Psychiatric, psychosocial, and cognitive functioning of female adolescents with ADHD. *J Am Acad Child Adolesc Psychiatry* 2001, **40**(5):530–540.
15. Kopp S, Kelly KB, Gillberg C: Girls with social and/or attention deficits: a descriptive study of 100 clinic attenders. *J Atten Disord* 2010, **14**(2):167–181.
16. Biederman J, Petty CR, Fried R, Doyle AE, Spencer T, Seidman LJ, Gross L, Poetzl K, Faraone SV: Stability of executive function deficits into young adult years: a prospective longitudinal follow-up study of grown up males with ADHD. *Acta Psychiatr Scand* 2007, **116**(2):129–136.
17. Barkley RA: Attention-deficit/hyperactivity disorder, self-regulation, and time: toward a more comprehensive theory. *J Dev Behav Pediatr* 1997, **18**(4):271–279.
18. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF: Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol Psychiatry* 2005, **57**(11):1336–1346.
19. Castellanos FX, Sonuga-Barke EJ, Milham MP, Tannock R: Characterizing cognition in ADHD: beyond executive dysfunction. *Trends Cogn Sci* 2006, **10**(3):117–123.
20. Seidman LJ, Biederman J, Monuteaux MC, Valera E, Doyle AE, Faraone SV: Impact of gender and age on executive functioning: do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Dev Neuropsychol* 2005, **27**(1):79–105.
21. Jonsdottir S, Bouma A, Sergeant JA, Scherder EJ: Relationships between neuropsychological measures of executive function and behavioral measures of ADHD symptoms and comorbid behavior. *Arch Clin Neuropsychol* 2006, **21**(5):383–394.
22. Egeland J: Frequency of attention deficit in first-episode schizophrenia compared to ADHD. *Appl Neuropsychol* 2010, **17**(2):125–134.
23. Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJ: Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biol Psychiatry* 2005, **57**(11):1224–1230.
24. McCue M, Pramuka M: Functional assessment. In *Rehabilitation*. Edited by Goldstein G, Beers S. New York: Plenum; 1998:113–129.
25. Gioia GA, Isquith PK, Guy SC, Kenworthy L: Behavior rating inventory of executive function. *Child Neuropsychol* 2000, **6**(3):235–238.
26. Anderson VA, Anderson P, Northam E, Jacobs R, Mikiewicz O: Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychol* 2002, **8**(4):231–240.
27. Bodnar LE, Prahme MC, Cutting LE, Denckla MB, Mahone EM: Construct validity of parent ratings of inhibitory control. *Child Neuropsychol* 2007, **13**(4):345–362.
28. Toplak ME, Bucciarelli SM, Jain U, Tannock R: Executive functions: performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychol* 2009, **15**(1):53–72.
29. McCandless S, O'Laughlin L: The clinical utility of the Behavior Rating Inventory of Executive Function (BRIEF) in the diagnosis of ADHD. *J Atten Disord* 2007, **10**(4):381–389.
30. Huizinga M, Smidts DP: Age-related changes in executive function: a normative study with the Dutch version of the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychol* 2011, **17**(1):51–66.
31. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N: Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 1997, **36**(7):980–988.
32. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R: *ADHD rating scale – IV. Checklists, norms & clinical interpretation*. New York: Guilford; 1998.
33. Achenbach TM, Rescorla LA: *Manual for the ASEBA School-Age Forms & Profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2001.
34. Wechsler D: *Wechsler abbreviated scale of intelligence. Norwegian version*. Stockholm: The Psychological Corporation; 1999.
35. Ivanova MY, Dobrea A, Dopfner M, Erol N, Fombonne E, Fonseca AC, Frigiero A, Grietens H, Hannesdottir H, Kanbayashi Y, et al: Testing the 8-syndrome structure of the child behavior checklist in 30 societies. *J Clin Child Adolesc Psychol* 2007, **36**(3):405–417.

36. Novik T: **Validity of the child behaviour checklist in a Norwegian sample.** *Eur Child Adolesc Psychiatry* 1999, **8**(4):247-254.
37. Novik T: **Child behavior checklist item scores in Norwegian children.** *Eur Child Adolesc Psychiatry* 2000, **9**(1):54-60.
38. Reynolds CR, Richmond BO: *Revised Children's Manifest Anxiety Scale (RCMAS)*. Manual. CA: Western Psychological Services; 1985.
39. Pela OA, Reynolds CR: **Cross-cultural application of the revised-Children's manifest anxiety scale: normative and reliability data for Nigerian primary school children.** *Psychol Rep* 1982, **51**(3):1135-1138.
40. Reynolds CR: **Long-term stability of scores on the revised-children's manifest anxiety scale.** *Percept Mot Skills* 1981, **53**(3):702-702.
41. Reynolds CR, Paget KD: **Factor analysis of the revised children's manifest anxiety scale for blacks, whites, males, and females with a national normative sample.** *J Consult Clin Psychol* 1981, **49**(3):352-359.
42. Scholwinski E, Reynolds CR: **Dimensions of anxiety among high IQ children.** *Gifted Child Quart* 1985, **29**(3):125-130.
43. Reynolds CR: **Concurrent validity of "what I think and feel:" the revised children's manifest anxiety scale.** *J Consult Clin Psychol* 1980, **48**(6):774-775.
44. Reynolds CR: **Convergent and divergent validity of the revised children's manifest anxiety scale.** *Educ Psychol Meas* 1982, **42**(4):1205-1212.
45. Turgeon L, Chartrand E: **Reliability and validity of the revised children's manifest anxiety scale in a french-canadian sample.** *Psychol Assess* 2003, **15**(3):378-383.
46. Boehnke K, Sillbereisen RK, Reynolds CR, Richmond BO: **What I think and feel: German experience with the revised form of the children's manifest anxiety scale.** *Personal Individ Differ* 1986, **7**(4):553-560.
47. Ferrando PJ: **Factorial structure of the revised children manifest anxiety scale in a Spanish sample: relations with Eysenck personality dimensions.** *Personal Individ Differ* 1994, **16**(5):693-699.
48. Spielberger CD, Edwards CD: *State-trait Anxiety Inventory for Children: STAIC: How I Feel Questionnaire*. Professional Manual. CA: Mind Garden; 1973.
49. Seligman LD, Ollendick TH, Langley AK, Baldacci HB: **The utility of measures of child and adolescent anxiety: a meta-analytic review of the revised children's manifest anxiety scale, the state-trait anxiety inventory for children, and the child behavior checklist.** *J Clin Child Adolesc Psychol* 2004, **33**(3):557-565.
50. Angold A, Costello EJ, Messer SC, Pickles A, Winder F, Silver D: **The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents.** *Int J Methods Psychiatr Res* 1995, **5**:1-12.
51. Costello EJ, Angold A: **Scales to assess child and adolescent depression: checklists, screens, and nets.** *J Am Acad Child Adolesc Psychiatry* 1988, **27**(6):726-737.
52. Costello EJ, Benjamin R, Angold A, Silver D: **Mood variability in adolescents: a study of depressed, nondepressed and comorbid patients.** *J Affect Disord* 1991, **23**(4):199-212.
53. Kovacs M: *The Children's Depression Inventory: A self-rated depression scale for school-aged youngsters*. Pittsburgh, PA: University of Pittsburgh School of Medicine; 1983.
54. Wechsler D: *Wechsler Intelligence Scale for Children-Fourth edition. Norwegian version*. Stockholm: The Psychological Corporation; 2004.
55. Stroop JR: **Studies of interference in serial verbal reactions.** *J Exp Psychol* 1935, **18**:643-662.
56. Delis D, Kaplan E, Kramer J: *Delis-Kaplan Executive Function System (D-KEFS). Norwegian version*. Stockholm: Pearson Assessment; 2001.
57. Fallmyr O, Egeland J: **Psychometric properties of the Norwegian version of BRIEF - for children from 5 to 18 years old.** *Norwegian Psychol Ass* 2011, **48**:339-343.
58. Breiman L: **Statistical modeling: the two cultures.** *Stat Science* 2001, **16**(3):199-231.
59. **Random forests. Version 4.0.** <http://www.stat.berkeley.edu/users/breiman/RandomForests/>.
60. Breiman: **Random forests.** *Mach Learn* 2001, **45**:5-32.
61. Liaw A, Wiener M: **Classification and regression by randomForest.** *R News* 2002, **2**(3):18-22.
62. Kuhn M: **Building predictive models in R using the caret package.** *J Stat Soft* 2008, **28**(5):1-26.
63. Novik TS, Hervas A, Ralston SJ, Dalsgaard S, Rodrigues Pereira R, Lorenzo MJ: **Influence of gender on attention-deficit/hyperactivity disorder in Europe-ADORE.** *Eur Child Adolesc Psychiatry* 2006, **15**(Suppl 1):15-24.
64. Graetz B, Sawyer M, Baghurst P: **Gender differences among children with DSM-IV ADHD in Australia.** *J Am Acad Child Adolesc Psychiatry* 2005, **44**(2):159-168.
65. Egger HL, Costello E, Erkanli A, Angold A: **Somatic complaints and psychopathology in children and adolescents: stomach aches, musculoskeletal pains, and headaches.** *J Am Acad Child Adolesc Psychiatry* 1999, **38**(7):852-860.
66. Egger HL, Angold A, Costello E: **Headaches and psychopathology in children and adolescents.** *J Am Acad Child Adolesc Psychiatry* 1998, **37**(9):951-958.
67. Rucklidge JJ, Tannock R: **Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender.** *J Child Psychol Psychiatry* 2002, **43**(8):988-1003.
68. Newcorn JH, Halperin JM, Jensen PS, Abikoff HB, Arnold LE, Cantwell DP, Conners CK, Elliott GR, Epstein JN, Greenhill LL: **Symptom profiles in children with ADHD: effects of comorbidity and gender.** *J Am Acad Child Adolesc Psychiatry* 2001, **40**(2):137-146.
69. Isquith P, Gioia G: **BRIEF predictions of ADHD: clinical utility of the Behavior Rating Inventory of Executive Function for detecting ADHD subtypes in children.** *Arch Clin Neuropsychol* 2000, **15**(8):780-781.
70. Weissman MM, Wickramaratne P, Warner V, John K, Prusoff BA, Merikangas KR, Gammon GD: **Assessing psychiatric disorders in children. Discrepancies between mothers' and children's reports.** *Arch Gen Psychiatry* 1987, **44**(8):747-753.
71. Lobovits DA, Handal PJ: **Childhood depression: prevalence using DSM-III criteria and validity of parent and child depression scales.** *J Pediatr Psychol* 1985, **10**(1):45-54.
72. Kashani JH, Orvaschel H, Burk JP, Reid JC: **Informant variance: the issue of parent-child disagreement.** *J Am Acad Child Psychiatry* 1985, **24**(4):437-441.
73. Rothen S, Vandeleur CL, Lustenberger Y, Jeanpretre N, Ayer E, Gamma F, Halfon O, Fornerod D, Ferrero F, Preisig M: **Parent-child agreement and prevalence estimates of diagnoses in childhood: direct interview versus family history method.** *Int J Methods Psychiatr Res* 2009, **18**(2):96-109.
74. Grills AE, Ollendick TH: **Issues in parent-child agreement: the case of structured diagnostic interviews.** *Clin Child Fam Psychol Rev* 2002, **5**(1):57-83.
75. Herjanic B, Reich W: **Development of a structured psychiatric interview for children: agreement between child and parent on individual symptoms.** *J Abnorm Child Psychol* 1982, **10**(3):307-324.
76. Herjanic B, Herjanic M, Brown F, Wheat T: **Are children reliable reporters?** *J Abnorm Child Psychol* 1975, **3**(1):41-48.
77. Faraone SV, Biederman J, Milberger S: **How reliable are maternal reports of their children's psychopathology? One-year recall of psychiatric diagnoses of ADHD children.** *J Am Acad Child Adolesc Psychiatry* 1995, **34**(8):1001-1008.
78. Gross-Tsur V, Goldzweig G, Landau YE, Berger I, Shmueli D, Shalev RS: **The impact of sex and subtypes on cognitive and psychosocial aspects of ADHD.** *Dev Med Child Neurol* 2006, **48**(11):901-905.
79. Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ: **Random forests for classification in ecology.** *Ecology* 2007, **88**(11):2783-2792.
80. Lambek R, Trillingsgaard A, Kadesjo B, Damm D, Thomsen PH: **Gender differences on the five to fifteen questionnaire in a non-referred sample with inattention and hyperactivity-impulsivity and a clinic-referred sample with hyperkinetic disorder.** *Scand J Psychol* 2010, **51**(6):540-547.

doi:10.1186/1471-244X-13-298

**Cite this article as:** Skogli et al.: ADHD in girls and boys - gender differences in co-existing symptoms and executive function measures. *BMC Psychiatry* 2013 **13**:298.



