J Abnorm Child Psychol (2013) 41:457–471
DOI 10.1007/s10802-012-9684-x

Executive Functions in Preschool Children with Externalizing Behavior Problems: A Meta-Analysis

Kim Schoemaker · Hanna Mulder · Maja Deković · Walter Matthys

Published online: 9 October 2012

Abstract Deficits in executive functions (EF) have been found in school-age children and adolescents with externalizing behavior disorders. Present meta-analysis was carried out to determine whether these EF impairments can also be found in preschool children with externalizing behavior problems. Twenty-two studies were included with a total of 4021 children. Four separate meta-analyses were conducted, concerning overall EF, working memory, inhibition and cognitive flexibility. A medium correlation effect size was obtained for overall EF (ESzr=0.22) and for inhibition (0.24), whereas a small effect size was found for working memory (0.17) and for cognitive flexibility (0.13). Moderator analyses revealed a stronger effect for older preschoolers compared to younger preschoolers, and for children from referred samples compared to community samples. These results show that EF, especially inhibition, is related to externalizing behavior problems already in preschool years.

Electronic supplementary material The online version of this article (doi:10.1007/s10802-012-9684-x) contains supplementary material, which is available to authorized users.

First submission to Journal of Abnormal Child Psychology: 25th of August 2011

K. Schoemaker (✉) · M. Deković · W. Matthys
Department of Child and Adolescent Studies, Utrecht University, P.O. Box 80140, 3508 TC Utrecht, the Netherlands
e-mail: k.schoemaker@umcutrecht.nl

K. Schoemaker · W. Matthys
Department of Child and Adolescent Psychiatry and Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, the Netherlands

H. Mulder
Department of Pedagogical and Educational Sciences, Utrecht University, Utrecht, The Netherlands

Keywords Executive Function · Externalizing behavior problems · Preschoolers · Meta-analysis

Although different definitions of executive functions (EF) exist, most authors agree that EF are the directing cognitive processes that enable purposeful and goal-directed behavior (Anderson 2002; Oosterlaan et al. 1998), i.e., the explicit control of thought, emotion and action (Séguin and Zelazo 2005). EF include mental processes such as planning, working memory, inhibition of inappropriate responses, flexibility in adaptation to environmental changes, and decision making (Nigg 2006). These functions serve to optimize behavior in changing environments. In their integrative framework, Miyake and colleagues (2000) proposed that in adulthood EF is a unitary construct with three partly dissociable components: working memory, inhibition and set shifting. Although the structure of EF that reflects both unity and diversity appears to be applicable from middle childhood onward (Wiebe et al. 2011), the structure during early childhood is not yet clearly defined.

Garon et al. (2008) reviewed the development of EF in normally developing preschoolers, based on the integrative framework of Miyake et al. (2000). They proposed a model in which each EF component is built upon earlier developing functions in the first years of life. In their model, the first developing EF component is working memory (keeping information in mind over a delay and in some tasks manipulating that information), followed by inhibition (withholding or delay of prepotent response). Both skills are integrated in set shifting, also referred to as cognitive flexibility (learning a rule and subsequently shifting to a new rule).

Besides this theoretical model, factor analytic studies have been conducted to empirically test the structure of EF in young children. Earlier studies identified three or four separable EF components in younger children using exploratory factor analyses (Hughes 1998; Espy et al. 1999; Murray and Kochanska 2002). More recently, however, studies used confirmatory factor analysis and found only one unitary executive construct in preschool children (Hughes et al. 2010; Wiebe et al. 2008, 2011; Willoughby et al. 2010). It must be noted that the latter three studies also examined a two-factor model (inhibition and working memory) which fitted the data equally well, but was rejected in favor of the one-factor model for reasons of parsimony. Thus, although the structure of EF in preschoolers still needs further investigation, recent studies indicate that EF may be characterized by a single factor structure in the preschool period and become increasingly differentiated with age (Hughes 2011; Wiebe et al. 2011).

Over the last decade, there has been an increasing interest in EF of children with hyperactive, impulsive and aggressive behavior. Research on EF in children with externalizing behavior disorders, however, has focused mainly on school-age children and adolescents. First, regarding working memory, in meta-analyses concerning attention deficit hyperactivity disorder (ADHD) medium to large effect sizes were found, indicating that working memory function is impaired in school-age children and adolescents with ADHD (Martinussen et al. 2005; Willcutt et al. 2005). In contrast, there are inconsistent results regarding working memory performance in adolescents with the two disruptive behavior disorders (DBD), i.e., oppositional defiant disorder (ODD) and conduct disorder (CD). On the one hand, Séguin and colleagues (1995, 1999) found a working memory impairment in physical aggressive children, also after controlling for ADHD. On the other hand, two more recent studies (Oosterlaan et al. 2005; Van Goozen et al. 2004) reported no differences between the ODD and normal control group.

Second, there is clear and convincing evidence of an inhibition deficit in school-age children and adolescents with ADHD with medium to large effect sizes, as has been repeatedly shown in meta-analyses (Alderson et al. 2007; Pennington and Ozonoff 1996; Willcutt et al. 2005). In a meta-analysis by Oosterlaan et al. (1998), however, deficits in inhibition were not uniquely associated with ADHD, but also with DBD. Other studies have shown deficiencies in inhibition with DBD, specifically when motivational processes, i.e., reward and punishment, are involved (Fairchild et al. 2009; Schutter et al. 2011).

Third, regarding cognitive flexibility, in meta-analyses concerning ADHD, medium effect sizes are reported, showing school-age children and adolescents with ADHD are at a disadvantage in this area (Pennington and Ozonoff 1996; Willcutt et al. 2005). Small effect sizes were found in the relation between cognitive flexibility and antisocial behavior (Morgan and Lilienfeld 2000). However, the possible role of ADHD co-morbidity was not examined in this study, although another meta-analysis revealed that EF deficits in CD are likely due to the presence of co-morbid ADHD (Pennington and Ozonoff 1996). In sum, clear EF deficits have been documented in school-age children and adolescents with ADHD, but the results regarding DBD are less consistent.

Although this research provides valuable information on the role of EF in school-age children and adolescents with ADHD and DBD, it is important to examine whether deficits in EF may already be observed in preschool children with externalizing behavior problems, as chronic patterns of hyperactivity and behavior problems can already be identified in the preschool years (Shaw et al. 2005). Moreover, programs to train EF skills in young children have been developed and proved to increase EF performance (Diamond and Lee 2011). Thus, if EF deficits occur already at preschool age in children with externalizing behavior problems, these children may benefit particularly from such training. Since EF is related to later academic performance, improvement in EF may result in better academic performance (Diamond and Lee 2011). Recently, the first meta-analysis was published in which neuropsychological deficits in preschoolers at risk for ADHD were examined (Pauli-Pott and Becker 2011). A small effect size was found for working memory and a medium to large effect size for inhibition tasks in children with ADHD symptoms in comparison to typically developing preschoolers. A conclusion regarding cognitive flexibility was not drawn, as cognitive flexibility tasks had been included only in a limited number of studies.

The present meta-analysis extends this work by investigating externalizing behavior problems in preschool children from a broader perspective, i.e., in addition to young children with ADHD symptoms, we also included studies in young children with aggressive and DBD symptoms. Moreover, stricter inclusion criteria were applied (e.g., definition of preschool age) and recent studies were included. Furthermore, in the current meta-analysis we included not only children with formal diagnoses, but also children with symptoms of ADHD and DBD. In many older studies, externalizing behavior problems in preschool children were viewed from a general perspective, e.g., 'hard to manage' children (Campbell et al. 1994; Hughes et al. 1998) as clinicians were reluctant to diagnose young children with a psychiatric disorder. Only recently, instruments have become available to differentiate between normal development and psychopathology at preschool age (e.g., Kiddie Disruptive Behavior Schedule, Keenan et al. 2007; Preschool Age Psychiatric Assessment, Egger and Angold 2004). Although our initial aim was to examine the relation between EF performance and ADHD or DBD separately, there were unfortunately too few studies on DBD to examine this relation. However, as there is a correlation of 0.79 between symptoms of DBD and ADHD hyperactivity in preschoolers (Sterba et al. 2007), it is presumably acceptable to examine ADHD and DBD combined.

The main aim of the current meta-analysis was to get specific insight in the relation between EF performance and externalizing behavior problems in preschool children. As there is an ongoing debate regarding the structure of EF in such young children and there are indications that EF may not be crystallized yet at this age, we studied EF at two levels; first overall EF, and second, the three EF components (working memory, inhibition, and cognitive flexibility) separately. Thus, the resulting two research questions are: (1) To what extent do preschool children with externalizing behavior problems exhibit overall EF impairments? (2) To what extent are the three EF components (working memory, inhibition, cognitive flexibility) related to externalizing behavior problems in preschool children? We hypothesize that possibly a smaller effect size for cognitive flexibility than for working memory and inhibition will be found, because cognitive flexibility develops later according to the developmental hierarchical model of Garon et al. (2008). Therefore, cognitive flexibility impairments might be less pronounced at this age.

The second aim of this meta-analysis was to examine whether there are factors that may moderate the relationship between EF and externalizing behavior problems. The strength of the association might depend in part on sample features (child age and gender) or sampling method. First, age may affect the relation between EF and behavior problems. Several studies confirm that EF performance in typically developing children improves rapidly with age during the preschool period (Carlson 2005; Hughes et al. 2010; Wiebe et al. 2008), with a developmental spurt between three and five years (Garon et al. 2008). Thus, the relation between EF and externalizing behavior problems is expected to be less clear at the beginning of this developmental process. The third research question therefore is: Is there a difference in the strength of the relation for older preschool children (4 ½ to 6 years) compared to younger (3 to 4 ½ years) preschool children?

Second, the strength of the relationship between EF and externalizing behavior problems may be influenced by the severity of the behavior problems. In community samples behavior problems are less distinct compared to referred samples. Assuming that the contribution of neurobiological factors is larger compared to environmental factors when the behavior problems are more severe, the strength of the association between EF and externalizing behavior problems will be stronger in referred samples in contrast to community samples. The results of a meta-analysis in older children with ADHD (Willcutt et al. 2005) showed that effect sizes on EF measures were slightly smaller in community samples compared to clinical samples and concluded that weaknesses in EF were not restricted to clinical samples but were also present in the general population. Thus, the fourth research question is: Is the magnitude of the effect sizes found for children from referred samples different from the magnitude of the effect sizes for children from community samples?

Third, the relation between EF and behavior problems may be affected by the child's gender. Girls could differ from boys in their EF performance due to their more rapid developmental maturation (Keenan and Shaw 1997). However, little is known about possible gender differences in the relation between externalizing behavior problems and EF in the preschool age. Raaijmakers et al. (2008) found that there was a stronger relation between aggression and EF in boys compared to girls, whereas Thorell and Wåhlstedt (2006) found no gender difference in EF performance in children with ODD/ADHD symptoms. The fifth research question is: Does gender play a role in the relationship between externalizing behavior problems and EF?

Method

Retrieval of Studies

A systematic computer search was performed in *Pubmed, Web of Science and PsychInfo*. The following combinations of keywords were used: Executive function (i.e., EF, working memory, inhibition, cognitive flexibility, neuropsychology), behavior problems (i.e., attention problems, attention deficit, hyperactivity, ADHD, oppositional, aggressive, externalizing, ODD, conduct problems, CD), and preschool (i.e., early childhood, young children). The reference lists of the retrieved articles were examined. Studies published before August 2011 were included. We limited our search to publications in English and in peer-reviewed journals. It should be noted that unpublished studies were excluded from this meta-analysis. Although limiting a meta-analysis to published documents introduces the potential for bias in favor of significant results, there is also a possibility of introducing a bias when searching for unpublished papers. It is not possible to systematically search for unpublished papers in a way that it is replicable. Moreover, there is no method available to calculate the magnitude of this bias (see also Kaminski et al. 2008). For this reason, we chose the standard and accepted method of computing the fail-safe number, rather than including unpublished studies with the risk of introducing an unknown source of bias.

Inclusion Criteria

To be included studies had to meet the following criteria: (a) the study included children with externalizing behavior problems (clinical diagnosis/ symptoms of ADHD, symptoms of ODD / CD, or aggressive or hard to manage children), (b) the mean age of the children in the study was between 3.0 and 6.0 years, and (c) EF tasks, aimed to measure working memory, inhibition or cognitive flexibility were administered. Studies were excluded which examined children with pediatric or neurological diseases.

In cases the article did not report sufficient information to permit calculation of the effect size, an attempt was made to contact the corresponding author for additional information. Data were obtained for four additional studies (Berwid et al. 2005; Raaijmakers et al. 2008; Tillman et al. 2008; Willoughby et al. 2010). In case of longitudinal studies, the first assessment was chosen. This resulted in choosing Campbell et al. (1994), Berlin and Bohlin (2002), Dalen et al. (2004), Thorell and Wåhlstedt (2006). Thus we excluded respectively Marakovitz and Campbell (1998), Berlin et al. (2003), Sonuga-Barke et al. (2002, 2003), Wåhlstedt et al. (2008). If within an article two assessments were reported on the same sample, the first assessment was included (Antshel and Nastasi 2008; Campbell et al. 1994). After correspondence with the authors it appeared that there was considerable overlap in sample between the studies of Berwid et al. (2005) and Marks et al. (2005). We included the study of Berwid, because the reported EF tasks resembled the other tasks included in this meta-analysis more.

The final sample included 22 studies (Table 1). Nineteen studies included one or more inhibition tasks, 13 studies included working memory tasks, and 5 studies included cognitive flexibility tasks. The included studies are marked with an asterisk in the reference list.

Coding the Studies

Each study was coded with a detailed coding scheme for recording sample, study and methodological characteristics. All these study characteristics were coded by the first author. In case of any doubt, the first author of the article was contacted.

EF Tasks

The EF tasks were categorized, based on the definition and categorization of Garon et al. (2008), as working memory, inhibition or cognitive flexibility tasks. The classification of the EF tasks was coded by the first and second author independently. Inter-rater agreement was 99 %. Although some EF tasks give multiple scores, it was often not possible to choose between different dependent variables from the same task, as most studies only reported a single dependent variable per task. This reflects the general heterogeneity in the assessment of EF which is characteristic of the field. Only in a few cases did we have the opportunity to choose between dependent variables. In those cases, the dependent variable was chosen which was most frequently used in the other included studies reporting on the same task (e.g., commission errors GoNoGo tasks). A table with the dependent variables of the EF tasks is displayed in online resource 2.

Working memory tasks are tasks in which children have to keep information in mind over a delay and, in some tasks, have to update or manipulate that information. The following tasks were included: Digit/word span (n=5), Spatial memory (n=3), Word span backwards (n=2), Delayed alternation (n=2), six/nine boxes (n=2), Selective reminding task (n=1), Sentence repetition (n=1), Span like task (n=1), Noisy book (n=1), Narrative memory (n=1), delayed response (n=1), Dual request selective task (n=1), Multiple boxes test (n=1), and Picture learning (n=1).

Inhibition tasks require withholding or delay of a prepotent or automatic response or holding a rule in mind, responding according to this rule, and inhibiting a prepotent response. The following tasks were found in the literature: Stroop tasks (n=8), GoNoGo (n=7), NEPSY statue (n=5), Delay of gratification (n=3), Snack delay (n=2), Shape School-inhibit condition (n=3), Knock & tap (n=2), Stop-Signal task (n=1), Resistance to temptation (n=1), Delay aversion (n=1), Detour reaching box (n=1), Spatial conflict (n=1), Puppet says (n=1), Luria's handgame (n=1), Pencil tapping (n=1), and Tongue task (n=1).

Cognitive flexibility tasks require forming an arbitrary stimulus–response set in the first phase and shifting to a new stimulus–response set in the second phase, with attention to a new aspect of the same stimulus. The tasks included were block sorting, Colour form test, Item selection, Object Classification Task, Set shifting, and Shape School-switch condition (for all tasks, n=1).

Externalizing Behavior Problems

Externalizing behavior problems were defined differently across studies, with some studies using a categorical approach, whereas others used a dimensional approach. In the categorical approach, used in 10 studies, a child was classified in the externalizing behavior problem group or the control group, according to either a cut-off score on a questionnaire or a judgment on the basis of a combination of different instruments. In these studies, the externalizing behavior problem group was defined as (high risk for) ADHD, hard to manage, aggressive, ODD or DBD. In the dimensional approach, used in 12 studies, symptoms or items on a continuous scale were used as outcome measures for behavior problems, including symptoms of ADHD, hyperactivity, attention problems, ODD, CD and aggression.

To assess externalizing behavior problems, a variety of instruments, informants and sampling methods were used

Table 1 Studies included in the meta-analysis, with descriptors and ESzr for each task

Study	N	Externalizing behavior problems	Age M + range (years; months)	Gender (% boys)	Sampling method	Assessment	EF	EF Task	ESzr task
Antshel and Nastasi (2008)	31	ADHD	4;9	60	Referred	Quest: P	WM	Picture learning	0.07
	31	Control	4;9 R: all 4 years	59		Interv: P			
Berlin and Bohlin (2002)	151	Hyperactive & CD	5;2	48	Community	Quest.: T	Inh	GoNoGo	0.38
Berwid et al. (2005)	16	high risk ADHD	4;8	64	Community	Quest : P & T	Inh	GoNoGo	0.21
	42	low risk ADHD	4;8 R: 3;5–6;11	63			Inh	Day-Night Stroop	0.16
Brocki et al. (2007)	72	ADHD and/or ODD	5;5	83	1/3 Selected by psychologist	Quest: P & T	WM	Digit Span- forward	0.13
			R: 4;1–7;0				WM	Digit Span- backward	0.06
							WM	Spatial memory	0.07
							Inh	NEPSY Statue	0.47
							Inh	Knock/tap/DNST	0.42
							Inh	GoNoGo	0.37
Campbell et al. (1994)	69	Hard2Manage	3;10	100	1/4 selected by parents (several with diagnosis)	Quest: P & T	Inh	Resistance to temptation	0.36
	43	Control	3;11 R: 2;6–4; 6	100			Inh	Delay of gratification	0.19
Dalen et al. (2004)	19	ADHD	3;3	–	First selection by health visitor	Interv: P	Inh	Delay aversion	0.40
	19	Control	3;3				Inh	Delay of gratification	0.92
			R: all 3 years				Inh	Puppet says	0.42
							CF	Block sorting	0.32
Dennis and Brotman (2003)	37	Aggression	6;0 R: 4;5–7;6	51	Community (prevention trial)	Quest: P	Inh	Four inhibition tasks of effortful control battery	0.39
Espy et al. (2011)	135–194	ADHD and ODD symptoms	3;11	44	Community	Quest: P	WM	Digit Span	0.09
			R: 2;5–6;0				WM	Delayed alternation	0.11
							WM	Delayed response	0.20
							WM	Six boxes	0.03
							Inh	NEPSY Statue	0.13
							Inh	Shape School- inhibit	0.12
Hughes et al. (1998)	40	Hard2Manage	4;4	60	Community	Quest: P & T	WM	Noisy book	0.19
	40	Control	4;2	60			Inh	Detour reaching box	0.30
			R: 3;6–4;6				Inh	Luria's handgame	0.23
							CF	Card sorting	0.17
Mahone et al. (2005)	40	ADHD	4;10	85	Referred	Quest: P	WM	Multiple boxes task	0.22
	40	Control	4;11 R: 3;0–6;6			Interv: P	Inh	NEPSY statue	0.36

Table 1 (continued)

Study	N	Externalizing behavior problems	Age M + range (years; months)	Gender (% boys)	Sampling method	Assessment	EF	EF Task	ESzr task
Mariani and Barkley (1997)	24	ADHD	5;0	100	Referred	Quest:	WM	Digit span	0.35
	30	Control	5;1	100		P or T	WM	Selective Reminding	0.42
			R: 4;0–5;11			Interv: P	WM	Spatial memory	0.43
							CF	Color form test	0.23
Perner et al. (2002)	24	At risk ADHD	5;9	–	Community	Quest: T	Inh	NEPSY statue	0.36
	22	Control	5;7			Interv: P	Inh	Knock & Tap	0.21
			R: 4;9–6;6				Inh	GoNoGo	0.31
							WM	Digit span -backward	0.12
Raaijmakers et al. (2008)	82	Aggressive behavior	4;3	72	Community	Quest : P	WM	Digit Span	0.16
	99	Control	4;4	65			Inh	GoNoGo	0.23
			R: all 4 years				Inh	Day-Night Stroop	0.18
							Inh	Shape School-inhibit	0.24
							CF	Shape School-switch	0.22
							CF	Object classification task	0.04
Re et al. (2010)	23 23	ADHD Control	5;7 5;9	61 61	Community	Quest: P & T	WM	Dual Request Selective Task	0.59
Rezazadeh et al. (2011)	31	Hyperactivity	4;9 R: 3–7 years	100	Community	Quest: P	Inh	Day-Night Stroop	0.07
Schoemaker et al. (2012)	61	ADHD	4;7	80	Referred	Quest:P & T	WM	Delayed alternation	0.20
	33	DBD	4;4	82		Interv: P	WM	Nine boxes	0.20
	52	ADHD + DBD	4;6	83			Inh	GoNoGo	0.37
	56	Control	4;8	70			Inh	Snack Delay	0.47
			R: 3;6–5;6				Inh	Shape School-inhibit	0.41
Thorell and Wåhlstedt (2006)	17	ADHD	5;0	77	Community	Quest: T	WM	Digit Span	0.17
	13	ODD	R: 4;0–6;2	46			WM	Spatial memory	0.17
	27 81	ADHD + ODD Control		70 40			Inh	Day-Night Stroop	0.30
Tillman et al. (2008)	107	ADHD hyp/imp	5;1	63	Community	Quest: T	Inh	Stop signal task	0.26
			R: 4;0–6;1				Inh	Stroop	0.37
Von Stauffenberg and Campbell (2007)	776	Attention problems	4;6	49	Community	Quest: T	Inh	Delay of gratification	0.16
							Inh	Day-Night Stroop	0.08

Table 1 (continued)

Study	N	Externalizing behavior problems	Age M + range (years; months)	Gender (% boys)	Sampling method	Assessment	EF	EF Task	ESzr task
Willoughby et al. (2010)	440–873	ADHD	3;1	51	Community	Quest: P, T, RA	WM	Span task	0.13
							Inh	Silly Sound	0.06
							Inh	stroop GoNoGo	0.05
							Inh	Spatial Conflict	0.12
							CF	Item Selection	0.12
Willoughby et al. (2011)	738–757	ADHD,ODD, aggression symptoms	4;6	50	Community (Head Start)	Quest: T	Inh	Pencil tapping	0.12
			R: all 4 years				Inh	Snack delay	0.13
							Inh	Tongue Task	0.07
Youngwirth et al. (2007)	28	Hyperactive	4;9	61	Community	Quest: P	WM	Sentence repetition	0.20
	14	ODD	R: 4;0–5;7	43		Interv: P	WM	Narrative memory	0.21
	29	Hyp + ODD		72			Inh	NEPSY Statue	0.19
	123	Control		55					

ADHD attention deficit hyperactivity disorder, *CF* cognitive flexibility task, *DBD* disruptive behavior disorder, *EF* executive function, *Interv* interview, *Inh* inhibition task, *ODD* oppositional defiant disorder, *P* parent report, *R* age range, *RA* research assistant, *Selected* selected from a community sample, *T* teacher report, *WM* working memory task, *Quest* questionnaire

across studies. Many studies used a combination of measurements and informants to define externalizing behavior problems. Sixteen different instruments were used to assess externalizing behavior problems, including eleven questionnaires and five semi-structured interviews. Many studies (k=10) used multiple informants (parents and a teacher/health visitor/research assistant). Eight of those studies reported on behavior problem groups, which were distinguished on the basis of reports from all informants and in the other two studies correlations were reported separately for different informants. In 7 studies questionnaires were only filled out by parents and in 5 studies questionnaires filled out by teachers were included. A table with information regarding the specific instruments, subscales, informants and how multiple measures were integrated is available upon request from the first author.

Differences in sampling method across studies were also noted. The majority of studies used a community sample (k=15), four studies reported on a clinically referred sample, and three studies used a procedure where they 'over selected' children with externalizing behavior problems in a community sample. This selection was done by identification of a child psychologist (Brocki et al. 2007), an observation of a health visitor (Dalen et al. 2004), or parent-identified (with several children in treatment for ADHD/ODD; Campbell et al. 1994). The definition of the externalizing behavior problems, informants and sampling method of each study is included in Table 1.

Moderators

Four moderators were included: age, sampling method, gender distribution, and impact factor of the journal. For age, we chose to split the mean age of the studies in the middle of the preschool period, resulting in a younger (3 years-0 months to 4 years-6 months) and older (4 years-7 months to 6 years-0 months) group. The sampling method was divided in referred/selected samples and community samples. Gender was coded as percentage of boys in the sample and analyzed as a continuous variable. The impact factor of a journal is a frequently used indicator of study quality in meta-analyses (e.g., Prinzie et al. 2009). The impact factor of the journal (2010) was analyzed as a continuous moderator. Few studies reported on IQ or ethnicity, and therefore, the potential moderator effects of these variables could not be tested adequately with the present set of studies.

Data Analysis

Initially, a single effect size for each EF task within a study was calculated (see Table 1). Subsequently, when studies reported results on different types of externalizing behavior, we averaged effect sizes across behavior problems. For example, when results on EF tasks were reported separately for the ODD and ADHD group within one study, we calculated the effect size of each group and subsequently averaged the effect sizes. There were two levels of analyses, resulting in four separate meta-analyses. First, a meta-analysis of the overall EF was conducted. In this analysis, the weighted mean effect size across all EF tasks in a study was computed, because within some studies there was considerable variation in the number of children completing each task. Then, the weighted mean effect size across studies was computed as described below. Second, meta-analyses were conducted for the separate EF components, i.e., working memory, inhibition and cognitive flexibility. For these analyses, first, a weighted mean effect size for all tasks of the same component was computed within a study to be entered in the meta-analysis.

For each meta-analysis, a mean standardized effect size (random effects model) and the 95 % confidence interval (CI) were computed. In twelve studies externalizing behavior problems were studied on a continuum, for these studies a correlation between the EF performance and externalizing behavior problems was reported. The other ten studies compared different groups, and subsequently reported means and standard deviations for each group or a Chi-square value (Hughes et al. 1998). We chose to convert all data to an effect size correlation (ESzr), as group comparisons are based on a dichotomy of an underlying continuous distribution of behavioral problems which is the same as in studies reporting correlations. All data were converted to ESzr with the transformations of Lipsey and Wilson (2001).

The fail- safe number (FSN) was calculated to indicate the robustness of findings. The FSN is the minimum number of studies with null results that are needed to reduce significant results to non-significance. If the FSN exceeds the critical value, i.e., five times the number of studies plus 10, findings are considered robust (recommended by Rosenthal 1995). The FSN also indicates sensitivity of the findings to publication bias (i.e., the tendency that non-significant findings often remain unpublished).

Homogeneity of the effect size distribution was examined with the Q-statistic. A significant Q-test indicates heterogeneity, which assumes that differences across effect sizes are due to sources other than sampling error, such as different study characteristics (Lipsey and Wilson 2001). Moderator analyses were then conducted to try to explain heterogeneity across effect sizes. If the meta-analyses of one of the EF components (working memory, inhibition, cognitive flexibility) were heterogeneous, additional task specific meta-analyses were conducted.

Moderator analyses for the continuous moderator (i.e., percentage boys and impact factor) were analyzed with weighted regression analyses and categorical moderators (i.e., age and sampling method) were analyzed with an ANOVA-procedure (both macro's by Lipsey and Wilson 2001). An ANOVA yields two homogeneity estimates, the Qbetween and Qwithin. A significant value for Qbetween indicates that the effects sizes are significantly different across different categories of the moderator variable, whereas a significant value for Qwithin indicates that the effect sizes within a category of the moderator variable are heterogeneous (Lipsey and Wilson 2001). The moderator variables were not related to each other, with the exception of gender and sample type. The clinical samples included studies with a higher percentage of boys ($F=15.45$, $p=0.001$).

Statistical meta-analyses and moderator analyses were performed using SPSS 16.0. Outliers were not excluded, considering the low number of studies. Correlation coefficient effect sizes can be classified as small (0.10), medium (0.25) or large (0.40; Lipsey and Wilson 2001).

Results

Sample and Study Characteristics

The 22 studies included in the meta-analyses provided data on 4021 children, with sample sizes ranging from 31 to 873 (Table 1). The range of EF tasks included per study was 1 to 6. The mean age of children across studies was 57 months (range 37 to 72 months). All studies reported on gender composition, with the exception of two studies. The percentage of boys ranged from 44 % to 100 %. Eight studies reported on the ethnicity of the children. The percentage of Caucasian ranged from 0 % (Dennis and Brotman 2003) to 100 % (Campbell et al. 1994; Mariani and Barkley 1997). Twelve studies reported a general IQ score, with the mean IQ across these studies ranging from 94 to 114.

All studies were published in peer reviewed journals between 1992 and August 2011. Twelve studies were conducted in the USA, four in Sweden, two in the United Kingdom, The Netherlands and Canada, and one in Italy and Austria. In all studies, EF performance and behavior problems were assessed concurrently, except for the study of Berlin and Bohlin (2002) in which the behavior problems were measured 9 months after the EF assessment.

Effect Sizes for Overall EF

There was a significant relationship between the level of overall EF performance and externalizing behavior problems with a medium effect size of 0.22 ($p<0.001$; see Table 2). This analysis included all 22 studies and 69 EF assessments, and the effect sizes ranged from 0.07 to 0.59. Figure 1 shows the effect sizes and 95 % CI for each study of the overall EF. The FSN was 547, which far exceeds Rosenthal's (1995) critical value (i.e., 120). The results can thus be considered robust against the file drawer effect.

Table 2 Effect sizes for overall EF and for separate EF components

	k	*N*	ESzr	95 % CI	Q - statistic
Overall EF					
Overall EF	22	4.021	0.22***	0.17–0.27	41.53**
EF factors					
Working memory	13	2.132	0.17***	0.12–0.23	15.06
Inhibition	19	3.795	0.24***	0.18–0.30	47.97***
Cognitive flexibility	5	1.198	0.13***	0.08–0.19	2.24
Inhibition tasks					
GoNoGo	8	1.238	0.26***	0.15–0.38	22.58**
Stroop	7	1.785	0.16***	0.08–0.25	14.55*
Delay with motivation	5	1.875	0.32***	0.16–0.48	37.16***
Delay without motivation	5	614	0.27***	0.14–0.40	8.62

k number of studies, *N* number of participants, *ESzr* mean correlation effect size, *CI* confidence interval

*$p<0.05$, **$p<0.01$, ***$p<0.001$

The Q-statistic testing effect size heterogeneity was significant, Q (21)=41.53, $p=0.005$, indicating the effect sizes may be drawn from different populations. The results of the categorical moderator analyses are presented in Table 3. Moderator analyses revealed a significantly larger effect size for studies with older preschoolers (4 1/2 to 6 years) relative to studies with younger preschoolers (3 to 4 1/2 years). Studies examining a referred/selected sample reported significantly larger effect sizes compared to studies from community samples. Regression analyses showed a significant effect of gender distribution ($\beta=0.44$, $p=0.029$), indicating that studies with a higher percentage of boys showed a higher effect size. There was no effect of the impact factor ($\beta=0.15$, $p=0.516$).

Effect Sizes for EF Components

Working Memory The meta-analysis of the working memory tasks included 13 studies and 23 assessments. A significant small effect size of 0.17 ($p<0.001$) was found. Effect sizes ranged from 0.07 to 0.59 (Figure 2 in Online Resource 1). The FSN of 183 exceeded Rosenthal's criteria of 75. The Q statistic was non-significant, Q (12)=15.06, $p=0.238$, indicating homogeneity.

Inhibition In the meta-analysis of inhibition, 19 studies and 39 assessments were included. Results indicated a medium significant effect size of 0.24 ($p<0.001$). Effect sizes ranged from 0.07 to 0.58 (see Figure 3 in Online Resource 1). Results were robust against the file drawer problem, as the

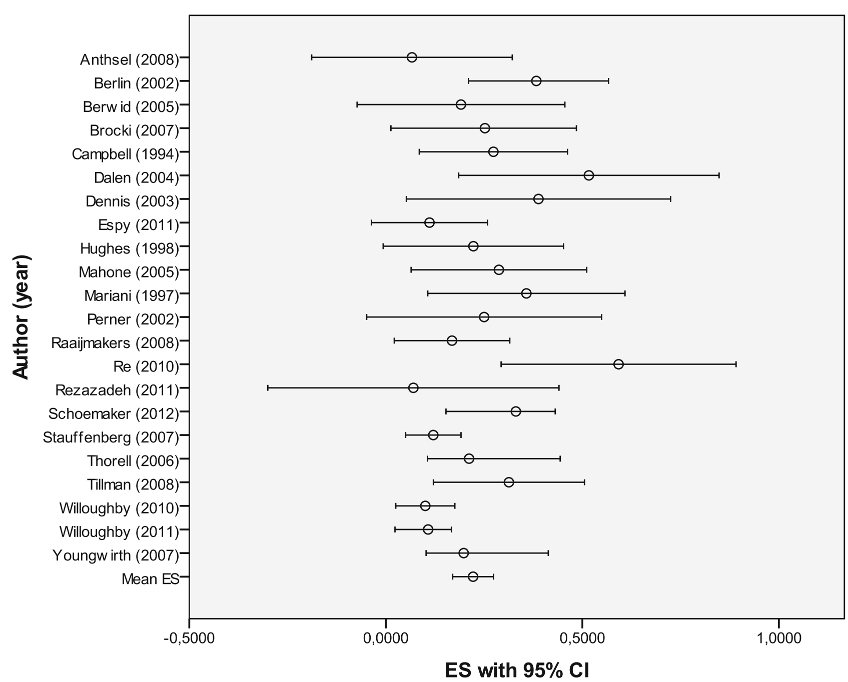

Fig. 1 Forest plot of overall EF tasks ($k=22$), mean effect size with 95 % CI

FSN of 413 exceeded by far the Rosenthal's (1995) criteria of 105.

The distribution of effect sizes was heterogeneous, Q (18)=47.97, $p<0.001$, indicating the effect sizes may be drawn from different populations. Moderator analyses showed a significant effect of age and sample, with larger effect sizes found in older preschool children and children from referred/selected samples. Regression analyses showed a significant effect of gender distribution (β=0.48 $p=0.026$), indicating that studies including a higher percentage of boys demonstrated a higher effect size. There was no effect of the impact factor (β=0.32, $p=0.194$).

Cognitive Flexibility The meta-analysis on cognitive flexibility included 5 studies and 6 assessments. Results indicated a small but significant effect size of 0.13 ($p<0.001$). Effect sizes ranged from 0.12 to 0.32 (Figure 4 in Online Resource 1). There is possibly a file drawer problem while the FSN was 34 and Rosenthal's criterion was 35, so these findings should be interpreted with precaution. The Q statistic was non-significant Q (4)=2.24, $p=0.692$, indicating homogeneity.

Additional Analyses: Effect Sizes for Inhibition Tasks

An additional factor influencing heterogeneity in the inhibition component might be the wide variety of tasks administrated. Therefore, additional meta-analyses were conducted on task level. Four categories of common inhibition tasks were created, i.e., GoNoGo tasks, Stroop tasks, Delay tasks with motivation and Delay tasks without motivation.

GoNoGo Tasks (Response Inhibition) In GoNoGo tasks children have to press a key when they see a target stimulus (Go condition) and inhibit that response when they see a non-target stimulus (NoGo condition). The meta-analysis for the GoNoGo tasks included 8 tasks, including the computerized GoNoGo tasks ($n=7$) and a modified Stop Signal Task (Tillman et al. 2008). In case of the GNG/CPT task (Berwid et al. 2005), we requested the data from the 5 (Go) vs. 1 (NoGo) ratio and included it here as a GoNoGo task. Results indicated a medium mean effect size 0.26 ($p<0.001$). The FSN was 55, which exceeds Rosenthal's critical value of 50. The results may thus be considered robust against publication bias. The Q-test indicated heterogeneity across samples, Q (7)=22.58, $p=0.002$.

Stroop Tasks (Interference Suppression) In the Stroop tasks children have to suppress an automatic response and activate a conflicting response. For example, in the day/night Stroop task, children must respond "night" to a picture of the sun and "day" to a picture of the moon. The Stroop analysis included 7 tasks, including the day/night, boy/girl and silly sound Stroop tasks. The mean effect size was small but significant, 0.16 ($p<0.001$). The FSN, 28, was under the critical value of Rosenthal, 45, indicating a possible file drawer problem. The Q-test indicated heterogeneity, Q (6)=14.55, $p=0.024$.

Delay Tasks (Response Delay with Motivation) In the delay tasks with tangible reward children had to stand (still) and wait for a signal before they could have a treat. The included tasks were Delay of gratification ($n=3$) and Snack delay ($n=2$). There was a medium significant effect size 0.32 ($p<0.001$). The FSN of 22 was under the critical value of Rosenthal, 35, indicating a possible file drawer problem. The Q-statistic was significant, Q (4)=37.16, $p<0.001$.

Table 3 Categorical moderator analysis

	$Q_{between}$	k	ESzr	95 % CI	Q_{within}
Overall EF ($k=22$)					
Age	20.90***				
Younger preschoolers (3.0–4.5)		7	0.12***	0.08–0.16	7.03
Older preschoolers (4.6–6.0)		15	0.28***	0.22–0.33	13.06
Sample	4.10*				
Community		15	0.18***	0.13–0.23	17.10
Referred/selected		7	0.29***	0.20–0.38	4.60
Inhibition ($k=19$)					
Age	19.18***				
Younger preschoolers (3.0–4.5)		7	0.14***	0.09–0.19	9.92
Older preschoolers (4.6–6.0)		12	0.31***	0.25–0.37	8.13
Sample	13.31***				
Community		14	0.18***	0.13–0.23	14.08
Referred / selected		5	0.39***	0.29–0.49	2.55

*$p<0.05$, **$p<0.01$, ***$p<0.001$

Delay Tasks (Response Delay Without Motivation) In the delay tasks without tangible reward children had to stand (still) and wait for a signal before they were allowed to move again. The analysis included the five NEPSY statue tasks. There was a medium significant effect size 0.27 ($p<0.001$). The FSN of 27 was just under the critical value of Rosenthal, 35, indicating a possible file drawer problem. The Q-test indicated homogeneity, Q (4)=8.62, $p=0.071$.

Additional Analyses: Separate Analyses for ADHD and DBD Symptoms

In order to investigate possible differences between ADHD and DBD symptoms, we conducted additional analyses, separately for studies that examined DBD symptoms and studies that examined ADHD symptoms. Studies were excluded if it was not possible to differentiate between ADHD and DBD (i.e., these studies included general categories of problem behavior, such as “hard to manage”). There were too few studies to conduct moderator and task specific analyses separately. There were 9 studies which examined DBD symptoms: 2 studies examining aggressive symptoms, 3 studies examined a DBD-only group (next to three other groups) and 4 studies reported a correlation with DBD symptoms (next to ADHD symptoms of the same group). Concerning the studies examining DBD symptoms, for overall EF we found an effect size of 0.19 ($p<0.001$, 95 % CI: 0.11–0.26, $k=9$), for working memory 0.15 ($p<0.001$, 95 % CI: 0.07–0.22, $k=6$), for inhibition 0.22 ($p<0.001$, 95 % CI: 0.13–0.31, $k=9$), and for cognitive flexibility 0.13 (one study). There were 18 studies which examined an ADHD-only group or ADHD symptoms. Analyses for ADHD symptoms show for overall EF an effect size of 0.21 ($p<0.001$, 95 % CI: 0.16–0.21, $k=18$), for working memory 0.17 ($p<0.001$, 95 % CI: 0.10–0.24, $k=11$), for inhibition 0.24 ($p<0.001$, 95 % CI: 0.17–0.31, $k=13$), and for cognitive flexibility 0.14 ($p<0.001$, 95 % CI: 0.06–0.22, $k=3$). In sum, the effect sizes for DBD were very similar to the effect sizes for ADHD (range of differences 0.01 to 0.02), providing further support for the decision to conduct the analyses for ADHD and DBD symptoms together. However, this result needs to be interpreted with caution due to the low number of studies.

Discussion

The purpose of the current meta-analysis was to determine the strength of the relationship between EF and externalizing behavior problems in preschool children. The unique contribution of this meta-analysis is the examination of the frequently co-occurring ADHD and DBD symptoms in preschoolers and the study of EF performance at different levels (i.e., overall, component, and task level) based on the latest insights in the structure of EF in preschoolers. The results indicate a medium mean effect size for the relationship between overall EF and externalizing behavior problems. However, results differed for the different EF components. A small effect size was found for working memory, a medium mean effect size was found for inhibition, and a small effect size for cognitive flexibility. The moderator analyses showed a stronger relation between EF and externalizing behavior problems for older preschool children (4 1/2–6 year) compared to younger preschool children (3–4 1/2 year) for the overall EF and inhibition component. There was a stronger relationship between EF and externalizing behavior problems for the children from referred/selected samples compared to community samples for the overall EF and inhibition component. Furthermore, there was a stronger relationship between EF and externalizing behavior problems in studies with a higher percentage of boys.

A medium effect size was found for *overall EF* (i.e., across all EF components and tasks).We chose to conduct a meta-analysis on overall EF given the findings of recent confirmatory factor analytic studies that EF may be best described as a single latent construct in very young children (Hughes et al. 2010; Wiebe, et al. 2008; Wiebe et al. 2011; Willoughby et al. 2010). However, the results of the present meta-analysis suggest that in preschool children with externalizing behavior problems there may not be just one EF factor, as there were differences in effect sizes across the different factors, with stronger impairments found for inhibition as compared to working memory and cognitive flexibility. If indeed there was only one overall EF factor, we would have expected the effect sizes of the different domains to be much more similar. Our finding is in line with the two-factor model of EF (inhibition and working memory) found in a recent study of preschoolers with ADHD and/or DBD (Schoemaker et al. 2012) and a three factor model in a sample with aggressive preschoolers (Raaijmakers et al. 2008). It might be that the EF factor structure is different in preschool children with externalizing behavior problems compared to typically developing children. However, the multi factor model is not necessarily at variance with the results of confirmatory factor analytic studies of typically developing preschool children (Wiebe et al. 2008, 2011; Willoughby et al. 2010), which reported that the two-factor model fitted equally well as the one-factor model. In fact, Wiebe et al. (2011) proposed that a multiple factor EF model might potentially be more ecologically valid than a single factor solution in preschoolers. The findings from the current meta-analysis seem to support this notion.

For *working memory*, a smaller effect size was found in the current study compared to meta-analyses of older children. Medium effect sizes were found for older children with ADHD, e.g., for spatial working memory and verbal working memory (d=0.63 and 0.55, respectively; Willcutt et al. 2005), compared to a small effect size in our study. The medium effect size concerning the *inhibition* factor and separate inhibition tasks found in this meta-analysis is consistent with the results of meta-analyses of older children. For example, medium effect sizes were found for children with ADHD for the Stop signal task (Cohen's d=0.61; Willcutt et al. 2005) and for children with antisocial behavior on the Stroop task (d=0.43; Morgan and Lilienfeld 2000). A similar pattern as working memory is observed for *cognitive flexibility*; we found a small effect size in contrast to meta-analyses of older children, which reported larger effect sizes. For example, on the Wisconsin Card Sorting Test a medium effect size was found for children with ADHD (d=0.46; Willcutt et al. 2005) and a small to medium effect size for children with antisocial behavior (d=0.24; Morgan and Lilienfeld 2000). The cognitive flexibility component showed the lowest effect size compared to the other EF components. There was even a non-overlapping 95 % confidence interval with inhibition. A possible explanation for this result may be that in the hierarchical model of Garon et al. (2008) cognitive flexibility is the last emerging EF component. As a result, it could be that at such young age children with cognitive flexibility impairments cannot be distinguished yet from typically developing children (see also Pauli-Pott and Becker 2011).

Age, sampling method and gender distribution were studied as moderators in the relationship between EF and externalizing behavior problems. The results of the present meta-analysis show that there is a stronger relation between overall EF performance (and specifically inhibition performance) and externalizing behavior problems in older compared to younger preschool children. The age effect in the present meta-analysis may be due to several reasons. First, there are concerns regarding validity and measurement error of EF tasks, particularly for the younger preschoolers, even though in recent years a shift has been observed away from simplifying adult tests and toward more developmentally appropriate tasks for assessing EF in young children (Carlson 2005; Espy et al. 1999; Hughes 1998). These child-friendly tasks are designed to minimize the complexity of instructions and responses. Despite these advantages, the individual variation in performance results not only from variation in EF ability but also from variation in non-executive abilities, i.e., language and motor skills (Wiebe et al. 2011). Second, among three year old children with medium to high levels of externalizing behavior problems, there is a subgroup of children who shows a decrease in these problem behaviors over the preschool period (Shaw et al. 2005). Possibly, the behavior problems of these three year old children are more associated with other risk factors (e.g., environmental) than with EF impairments, resulting in a weaker relation between EF and externalizing behavior problems. A longitudinal design is needed to test the hypothesis that EF impairment is a risk factor that plays a role in the stability of externalizing behavior problems in the preschool period.

Regarding sampling method, this meta-analysis clearly shows that EF impairments are not restricted to preschool children from clinical samples, but are also evident in preschool children from community samples. Small effect sizes were found in the community samples and medium to large effect sizes were found in the referred/selected samples. There was a significant difference between the sampling methods for the overall EF and inhibition domain. This result indicates a stronger relationship between EF performance and externalizing behavior problems when the behavior problems are more severe. Thus, the relation between EF (and especially inhibition task) performance and externalizing behavior problems may be nonlinear in nature. We assume that the role played by EF as a risk factor is less pronounced in community samples compared to referred samples.

Gender distribution of the sample appeared to have an effect on the relationship between EF and externalizing behavior problems. This result indicates a stronger relation between EF performance and externalizing behavior problems when studies included a higher percentage of boys. As it was not possible in this meta-analysis to directly compare EF performance between preschool boys and girls with behavior problems, further empirical studies are needed to examine this issue.

Recently, a meta-analysis was published investigating neuropsychological functioning in preschoolers with ADHD symptoms (Pauli-Pott and Becker 2011). The data used for the current meta-analysis differs in a number of ways from the data used by Pauli-Pott and Becker (2011). First, we not only included studies measuring ADHD symptoms, but we also included studies of children with aggressive and DBD symptoms. Second, we added seven more recent studies. Third, we strictly defined preschool age, therefore we limited ourselves to studies with a mean sample age until 6 years age, whereas Pauli-Pott and Becker (2011) included studies with a mean age until 7 years. Fourth, in contrast to the other meta-analysis, we included only the first of multiple assessments of the same sample to avoid dependency between data.

Even though additional studies and groups were included and stricter inclusion criteria were used, comparable results were found in the two meta-analyses, with slightly higher effect sizes found in the meta-analysis by Pauli-Pott and Becker (2011). However, the results of the current meta-analysis extend the findings of Pauli-Pott and Becker. First, additional evidence is provided on an overall EF and factor level instead of only on task level. Second, since the present meta-analysis includes EF data from preschool children with ADHD and DBD symptoms, this meta-analysis provides a broader perspective on EF impairments in preschoolers with externalizing behavior problems rather than specific for ADHD. This is important as in the preschool period ADHD and DBD symptoms are strongly associated (Sterba et al. 2007).

Finally, a number of limitations of this meta-analysis need to be considered. First, there was quite a broad range of operationalizations of EF (e.g., 16 different inhibition tasks) and externalizing behavior problems across studies. This could have influenced the strength and specificity of the results. Second, additional moderators could influence the strength of the relationship, e.g., Social Economic Status of the family, ethnicity, and IQ. Although these moderators were considered at the set-up of this meta-analysis, there was not enough information available in the studies to conduct these analyses. Third, in about half of the studies the sample size was relatively small. Fourth, studies were included in which the mean age of the children was between 3.0 and 6.0 years. Consequently, there were a number of studies in which some of the children were older than 6.0 years, therefore the finding should be interpreted with caution.

The findings of the present study have clinical implications. First, traditionally much attention has been given to environmental factors that affect young children with externalizing behavior problems (Campbell et al. 1994). The last years, however, more research has been conducted on child characteristics. This meta-analysis clearly shows that EF impairments can be identified in preschool children with externalizing behavior problems. Second, programs to train EF skills have been developed (Diamond and Lee 2011). Such programs have been shown to increase EF performance, especially children with the initially poorest executive functions benefit most from these programs. EF predicts later academic performance, so improvement in EF may lead to improvement of school readiness and academic achievement (Diamond and Lee 2011). Thus, early executive function training may reduce the achievement gaps later.

In conclusion, this meta-analysis shows that EF impairments in children with externalizing behavior problems can already be identified in the preschool period. However, more research needs to be undertaken to improve our understanding of the association between EF and externalizing behavior problems at this young age. An important limitation of the current literature is that little attention has been paid to differentiate EF deficits in DBD and in ADHD. Even though this is a challenge due to the high correlation between ADHD and DBD symptoms, future studies should assess EF performance for ADHD and DBD separately (e.g., Schoemaker et al. 2012; Youngwirth et al. 2007). Most importantly, longitudinal studies are needed to investigate the role of EF impairments in the stability of externalizing behavior problems.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. *Journal of Abnormal Child Psychology, 35*, 745–758.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*, 71–82.

*Antshel, K. M., & Nastasi, R. (2008). Metamemory development in preschool children with ADHD. *Journal of Applied Developmental Psychology, 29*, 403–411.

*Berlin, L., & Bohlin, G. (2002). Response inhibition, hyperactivity, and conduct problems among preschool children. *Journal of Clinical Child Psychology, 31*, 242–251.

Berlin, L., Bohlin, G., & Rydell, A. M. (2003). Relations between inhibition, executive function, and ADHD symptoms: a longitudinal study from age 5 to 8 1/2 years. *Child Neuropsychology, 9*, 255–266.

*Berwid, O. G., Curko Kera, E. A., Marks, D. J., Santra, A., Bender, H.A., & Halperin, J.M (2005). Sustained attention and response inhibition in young children at risk for attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry, 46*, 1219–1229.

*Brocki, K. C., Nyberg, L., Thorell, L. B., Bohlin, G. (2007). Early concurrent and longitudinal symptoms of ADHD and ODD: relations to different types of inhibitory control and working memory. *Journal of Child Psychology and Psychiatry, 48,* 1033–1041.

*Campbell, S. B., Pierce, E. W., March, C. L., Ewing, L. J., & Szumowski, E. K. (1994). Hard-to-manage preschool boys: symptomatic behavior across contexts and time. *Child Development, 65*, 836–851.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*, 595–616.

*Dalen, L., Sonuga-Barke, E. J. S., Hall, M., & Remington, B. (2004). Inhibitory deficits, delay aversion and preschool AD/HD: implications for the dual pathway model. *Neural Plasticity, 11*, 1–11.

*Dennis, T. A., & Brotman, L. M. (2003). Effortful control, attention, and aggressive behavior in preschoolers at risk for conduct problems. *Annals of the New York Academy of Sciences*, *1008,* 252–255.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*, 959–964.

Egger, H. L., & Angold, A. (2004). The preschool Age psychiatric assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. In R. Delcarmen-Wiggens & A. Carter (Eds.), *Handbook of infant, toddler, and preschool mental assessment* (pp. 223–243). New York: Oxford University Press.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., & Gilsky, M. L. (1999). Executive functioning in preschool children: performance on A-not-B and other delayed response format tasks. *Brain and Cognition, 41*, 178–199.

*Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A. C., & Moehr, M. J. (2011). Executive control and dimensions of problem behaviors in preschool children. *Journal of Child Psychology and Psychiatry, 52,* 33–46.

Fairchild, G., van Goozen, S. H. M., Stollery, S. J., Aitken, M. R. F., Savage, J., Moore, S. C., & Goodyer, I. M. (2009). Decision making and executive function in male adolescents with early-

onset or adolescence-onset conduct disorder and control subjects. *Biological Psychiatry, 66*, 162–168.
Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin, 134*, 31–60.
Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*, 233–253.
Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development, 20*, 251–271.
*Hughes, C., Dunn, J., & White, A. (1998). Trick or treat? Uneven understanding of mind and emotion and executive dysfunction in hard-to-manage preschoolers. *Journal of Child Psychology and Psychiatry, 39,* 981–994.
Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2010). Tracking executive function across the transition to school: a latent variable approach. *Developmental Neuropsychology, 35*, 20–36.
Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36*, 567–589.
Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin, 121*, 95–113.
Keenan, K., Wakschlag, L. S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 457–468.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage.
*Mahone, E. M., Pillion, J. P., Hoffman, J., Hiemenz, J. R., & Denckla, M. B. (2005). Construct validity of the auditory continues performance test for preschoolers. *Developmental Neuropsychology, 27*, 11–13.
Marakovitz, S. E., & Campbell, S. B. (1998). Inattention, impulsivity, and hyperactivity from preschool to school age: performance of hard-to-manage boys on laboratory measures. *Journal of Child Psychology and Psychiatry, 39*, 841–851.
*Mariani, M. A., & Barkley, R. A. (1997). Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. *Developmental Neuropsychology, 13*, 111–129.
Marks, D. J., Berwid, O. G., Santra, A., Kera, E. C., Cyrulnik, S. E., & Halperin, J. M. (2005). Neuropsychological correlates of ADHD symptoms in preschoolers. *Neurpsychology, 19*, 446–455.
Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 377–384.
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*, 49–100.
Morgan, A. B., & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review, 20*, 113–136.
Murray, K. T., & Kochanska, G. (2002). Effortful control: factor structure and relation to externalizing and internalizing behaviors. *Journal of Abnormal Child Psychology, 30*, 503–514.
Nigg, J. T. (2006). *What causes ADHD? Understanding what goes wrong and why*. New York: The Guilford Press.
Oosterlaan, J., Logan, G. D., & Sergeant, J. A. (1998). Response inhibition in AD/HD, CD, comorbid AD/HD + CD, anxious, and control children: a meta-analysis of studies with the stop task. *Journal of Child Psychology and Psychiatry, 39*, 411–425.
Oosterlaan, J., Scheres, A., & Sergeant, J. A. (2005). Which executive functioning deficits are associated with AD/HD, ODD/CD and comorbid AD/HD + ODD/CD? *Journal of Abnormal Child Psychology, 33*, 69–85.
Pauli-Pott, U., & Becker, K. (2011). Neuropsychological basic deficits in preschoolers at risk for ADHD: a meta-analysis. *Clinical Psychology Review, 31*, 626–637.
Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51–87.
*Perner, J., Kain W., & Barchfield, P. (2002). Executive control and higher-order theory of mind in children at risk of ADHD. *Infant and Child Development, 11*, 141–158.
Prinzie, P., Stams, G. J., Deković, M., Reijntjes, A. H. A., & Belsky, J. (2009). The relations between parents' Big Five personality factors and parenting: a meta-analytic review. *Journal of Personality and Social Psychology, 97*, 351–362.
*Raaijmakers, M. A. J., Smidts, D. P., Sergeant, J. A., Maassen, G. H., Posthumus, J. A., van Engeland, H., et al. (2008). Executive functions in preschool children with aggressive behavior: impairments in inhibitory control. *Journal of Abnormal Child Psychology, 36*, 1097–1107.
*Re, A., De Franchis, V., & Cornoldi, C. (2010). Working memory control deficit in kindergarten ADHD children. *Child Neuropsychology, 16*, 134–144.
*Rezazadeh, S. M., Wilding, J., & Cornish, K. (2011). The relationship between measures of cognitive attention and behavioral ratings of attention in typically developing children. *Child Neuropsychology, 17*, 197–208.
Rosenthal, R. (1995). *Meta-analytic procedures for social research*. Newbury Park: Sage.
*Schoemaker, K., Bunte, T., Wiebe, S. A., Espy, K. A., Deković, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry, 53,* 111–119.
Schutter, D. J. L. G., van Bokhoven, I., Vanderschuren, L. J. M. J., Lochman, J. E., & Matthys, W. (2011). Risky decision making in substance dependent adolescents with a disruptive behavior disorder. *Journal of Abnormal Child Psychology, 39*, 333–339.
Séguin, J. R., & Zelazo, P. D. (2005). In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Executive function in early physical aggression*. New York: Guilford.
Séguin, J. R., Pihl, R. O., Harden, P. W., Tremblay, R. E., & Boulerice, B. (1995). Cognitive and neuropsychological characteristics of physically aggressive boys. *Journal of Abnormal Psychology, 104*, 614–624.
Séguin, J. R., Boulerice, B., Harden, P. W., Tremblay, R. E., & Pihl, R. O. (1999). Executive functions and physical aggression after controlling for attention deficit hyperactivity disorder, general memory and IQ. *Journal of Child Psychology and Psychiatry, 40*, 1197–1208.
Shaw, D. S., Lacourse, E., & Nagin, D. S. (2005). Developmental trajectories of conduct problems and hyperactivity from ages 2 to 10. *Journal of Child Psychology and Psychiatry, 46*, 931–942.
Sonuga-Barke, E. J. S., Dalen, L., & Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 1335–1342.
Sonuga-Barke, E. J. S., Dalen, L., Daley, D., & Remington, B. (2002). Are planning, working memory, and inhibition associated with individual differences in preschool ADHD symptoms? *Developmental Neuropsychology, 21*, 255–272.
Sterba, S., Egger, H. L., & Angold, A. (2007). Diagnostic specificity and nonspecificity in the dimensions of preschool psychopathology. *Journal of Child Psychology and Psychiatry, 48*, 1005–1013.

*Thorell, L. B., & Wåhlstedt, C. (2006). Executive functioning deficits in relation to symptoms of ADHD and/or ODD in preschool children. *Infant and Child Development*, *15*, 503–518.

*Tillman, C. M., Thorell, L. B., Brocki, K. C., & Bohlin, G. (2008). Motor response inhibition and execution in the stop-signal task: development and relation to ADHD behaviors. *Child Neuropsychology, 14*, 42–59.

Van Goozen, S. H. M., Cohen-Kettenis, P. T., Snoek, H., Matthys, W., Swaab-Barneveld, H., & van Engeland, H. (2004). Executive functioning in children: a comparison of hospitalized ODD and ODD/ADHD children and normal controls. *Journal of Child Psychology and Psychiatry, 45*, 284–292.

*Von Stauffenberg, C., & Campbell, S. B. (2007). Predicting the early developmental course of symptoms of attention deficit hyperactivity disorder. *Journal of Applied Developmental Psychology, 28*, 536–552.

Wåhlstedt, C., Thorell, L. B., & Bohlin, G. (2008). ADHD symptoms and executive function impairment: early predictors of later behavioral problems. *Developmental Neuropsychology, 33*, 160–178.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44*, 575–587.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*, 436–452.

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*, 1336–1346.

*Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2010). The measurement of executive function at age 3 years: psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment, 22*, 306–317.

*Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental Neuropsychology, 36*, 162–180.

*Youngwirth, S. D., Harvey, A. H., Gates, E. C., Hashim, R. L., & Friedman-Weieneth, J. L. (2007). Neuropsychological abilities of preschool-aged children who display hyperactivity and/or oppositional-defiant behavior problems. *Child Neuropsychology, 13*, 422–443.