Psychological Medicine, 2005, **35**, 1097–1108.
doi:10.1017/S003329170500499X Printed in the United Kingdom

REVIEW ARTICLE

Executive functioning in adult ADHD: a meta-analytic review

A. MARIJE BOONSTRA[1]*, JAAP OOSTERLAAN[2], JOSEPH A. SERGEANT[2] AND JAN K. BUITELAAR[3]

[1]*Department of Psychology, Erasmus University, Rotterdam, The Netherlands*; [2]*Department of Clinical Neuropsychology, Vrije Universiteit, Amsterdam, The Netherlands*; [3]*Department of Psychiatry, University Medical Center St. Radboud, Nijmegen, The Netherlands*

ABSTRACT

Background. Several theoretical explanations of ADHD in children have focused on executive functioning as the main explanatory neuropsychological domain for the disorder. In order to establish if these theoretical accounts are supported by research data for adults with ADHD, we compared neuropsychological executive functioning and non-executive functioning between adults with ADHD and normal controls in a meta-analytic design.

Method. We compared 13 studies that (1) included at least one executive functioning measure, (2) compared the performance of an adult ADHD group with that of an adult normal control group, (3) provided sufficient information for calculation of effect sizes, and (4) used DSM-III-R or DSM-IV criteria to diagnose ADHD.

Results. We found medium effect sizes both in executive functioning areas [verbal fluency ($d=0{\cdot}62$), inhibition ($d=0{\cdot}64$ and $d=0{\cdot}89$), and set shifting ($d=0{\cdot}65$)] and in non-executive functioning domains [consistency of response ($d=0{\cdot}57$), word reading ($d=0{\cdot}60$) and color naming ($d=0{\cdot}62$)].

Conclusions. Neuropsychological difficulties in adult ADHD may not be confined to executive functioning. The field is in urgent need of better-designed executive functioning tests, methodological improvements, and direct comparisons with multiple clinical groups to answer questions of specificity.

INTRODUCTION

For many years, psychological research into attention deficit hyperactivity disorder (ADHD) has focused on attention problems as the core deficit (Douglas, 1999). More recently, some authors see the symptoms of ADHD as the consequence of disturbances in executive functioning (EF). Welsh & Pennington (1988) defined EF as follows: 'the ability to maintain an appropriate problem solving set for attainment of a future goal' (p. 201). Following this definition, Pennington & Ozonoff (1996) indicated five domains of EF: fluency (the ability to generate different solutions for a problem), planning (the ability to plan the steps needed to reach a solution for a problem), working memory (the ability to keep information online while performing), inhibition (the ability to inhibit or withhold one's actions), and set shifting (the ability to shift to another action or problem-solving set when necessary). Pennington & Ozonoff (1996) concluded that ADHD is associated with deficits in behavioral inhibition. In Barkley's (1997*b*) theory of ADHD, a core deficit in inhibition causes difficulties with many other EFs, such as working memory,

* Address for correspondence: A. Marije Boonstra, M.Sc., Erasmus University, Department of Psychology, Burgemeester Oudlaan 50, 3062 PA, Rotterdam, The Netherlands.
(Email: boonstra@fsw.eur.nl)

self-regulation, and motor control. Many researchers have indeed noted poorer performance on neuropsychological tasks designed to measure EF. Sergeant *et al.* (2002) reviewed studies using EF tasks in children with ADHD and related disorders. They reported clear evidence for EF deficits in ADHD in children, although they questioned the specificity of EF problems for this disorder, since many other childhood psychiatric disorders (e.g. oppositional defiant disorder, conduct disorder) are also related to deficits in EF. In a recent qualitative review, Woods *et al.* (2002) discussed studies in which EF measures were used with an adult ADHD sample. They concluded, 'that adults with ADHD demonstrate subtle impairments on select measures of attention and executive functions, auditory-verbal list learning, and complex information processing speed relative to normal controls' (p. 12). They further concluded 'The most prominent and reliable measures that differentiate adults with ADHD from healthy control samples were the various Stroop tasks, verbal letter fluency, auditory-verbal list learning, and continuous performance tests' (p. 28).

However much we commend the qualitative and narrative review of Woods *et al.* (2002), refinement of their conclusions can be found in a statistical or quantitative review of the literature. It is for this reason that we conducted the current meta-analytic review to quantitatively establish the difference between adults with ADHD and normal controls (NC) in EF. We compared studies using EF tests in a group of adults with ADHD and a group of NC adults. Since many of these tests also provide information on non-EF neuropsychological functions (e.g. speed of information processing, verbal memory) and since there are indications that not only EF is impaired in ADHD (e.g. Woods *et al.* 2002), we decided to also include non-EF variables from the EF tasks in our meta-analysis.

METHOD

Papers for consideration were identified through a literature search in PsychINFO, MEDLINE, and Current Contents from 1970 (around this time adult ADHD was first mentioned in the literature) through September 2003.

To be included in the analysis, studies had to meet the following criteria:

- Each study had to include at least one EF measure in one or more of five domains, as stated by Pennington & Ozonoff (1996).
- Studies had to compare the performance of an adult ADHD group (age above 18 years) with a group of NC participants.
- Sufficient information for calculation of effect sizes (ES) had to be available either directly from the paper, or through the contact author of the study.
- ADHD diagnoses had to be made according to either DSM-III-R or DSM-IV criteria.

We included only EF measures that had formerly been shown to rely on functioning of the frontal cortex, either in patient studies or by use of neuroimaging techniques. Further, an EF measure was only included in the study if at least four studies with an adult ADHD sample provided information on the same version of the test and on the same dependent variables, either directly in the paper or through contacting authors. Next to this criterion of four studies, both the total number of ADHD participants and the total number of NC participants in all studies had to exceed 50 for each dependent measure, in order to obtain enough power (0·80) to find significant results for at least medium ESs (Cohen, 1988).

EF measures

Controlled Oral Word Association (*COWAT*)

The COWAT (Spreen & Benton, 1977) is a test for verbal fluency. It assesses the capacity to produce different words starting with a specific letter within a specified time interval. The dependent variable used in this meta-analysis was the total number of correct words generated for three letters (F, A, and S, or C, F, and L) in 1 minute per letter.

Continuous Performance Test (*CPT*)

The version of the CPT used for our analyses is the Multi Health System Standard Task (Conners, 1995). The task requires participants to press the space bar as quickly as possible when they are presented with a letter on a computer screen. They have to do this for every letter except for the letter X, in which case they are to withhold their response. The most often reported (and therefore chosen for our analyses) dependent variables are: (1) mean reaction time for hits (hit RT; to measure the latency of the response execution process); (2) the standard error of the mean hit reaction time (SE hit RT; an indication of the consistency with which respondents can focus their attention); (3) the number of commission errors (COM), measuring inhibitive behavior (high error rates indicate poor inhibitive control); (4) the number of omission errors (OM; indicating poor vigilance); (5) attentiveness (d'; usually termed 'sensitivity' in signal detection theory), which is an indication of the ability to discriminate between targets (X) and non-targets (other letters); (6) Risk taking (β). This variable notifies a person's response tendency: higher values point to cautious response styles.

WAIS Digit Span (*DS*)

In the Wechsler Adult Intelligence Scale – Revised (WAIS-R; Wechsler, 1981) subtest DS, participants are required to repeat a series of digits read aloud by the experimenter. In DS-Forwards (DS-FW), the participant has to repeat the series in the same order it was read. This is a direct measure of verbal memory, with few EF connotations. In DS-Backwards (DS-BW) the series has to be repeated backwards. This manipulation requires working memory. Series of increasing difficulty level are presented. Dependent variables can be calculated separately for DS-FW and DS-BW by counting the number of correctly repeated series for each condition.

Stroop Color Word Test (*Stroop*)

In this measure of interference or mental inhibition (first developed by Stroop in 1935), a participant is shown three different cards. The first two cards require reading color names (card W) and naming colors (card C). The third card (color word: CW) is the actual interference card, which consists of color names, printed either in the denoted color (RED printed in red ink) or in a different color (RED printed in green ink). Participants are required to name the color of the ink rather than the name of the color. Often, the number of correctly named colors on card CW is chosen to represent interference. This is one of the dependent measures chosen in this meta-analysis. However, the validity of this variable as an indication of interference could be questioned, since performance on the first two cards may influence scores on the CW card. Hammes (1971) has, therefore, suggested correcting the score on the CW card for color-naming performance. We calculated this interference score with the raw mean data and included it as a second dependent variable in our analyses.

Trailmaking Test (*TMT*)

This test (Reitan & Wolfson, 1985) requires participants to connect series of circles. In part A (TMT-A), the circles contain numbers (1–25) and participants are instructed to connect them in counting order. This part requires serial information processing, visual scanning, and motor speed. Part B (TMT-B) contains circles with numbers and circles containing letters. The instruction is to connect the circles by alternating between numbers and letters (i.e. 1–A–2–B, etc.). TMT-B can be considered a measure of both working memory and interference control (inhibition). The dependent variables for both part A and B are the number of seconds needed to complete the sequence.

Calculation of ES and tests of homogeneity

All data were analyzed using the program Comprehensive Meta Analysis (Borenstein & Rothstein, 1999). We report Cohen's d (Cohen, 1988), which is defined as the difference between two means divided by standard deviation of either group. We corrected for sample size-bias with Hedges' formula (Hedges & Olkin, 1985). Our ES are, therefore, slightly more conservative than uncorrected ones, although differences between corrected and uncorrected indices are usually slight (Kulik & Kulik, 1989). The closer Cohen's d comes to zero, the smaller the difference between two groups. For each dependent variable, the ES from each study are combined into a grand mean estimate of the difference in performance between ADHD participants and NC participants. In accordance with Cohen (1988), we consider values between 0·2 and 0·5 as small, between 0·5 and 0·8 as medium, and above 0·8 as large.

In a meta-analysis, it is assumed that all ES are derived from a single population. The amount of variation (i.e. heterogeneity) within the established ES is reflected by the Q statistic

(Hedges & Olkin, 1985). If ES are homogeneous, this Q statistic will not exceed a critical value associated with an *a priori* established alpha level (in this study $p=0{\cdot}05$). If ES are not homogeneous, this could imply that other factors than chance and EF have influenced the results. An overview of these potential moderator variables will be provided in the Results section.

RESULTS

EF

We obtained data on five EF tests in 13 different studies that met our criteria for inclusion (see Table 1).

The results of the analyses of the EF measures are summarized in Table 2. Positive ES (Cohen's d) indicate a better performance for the NC group, while negative ES point toward an advantage for those with ADHD.

As can be concluded from the values of the Q statistic in Table 2, heterogeneity in ES was found for the COWAT, CPT risk taking, and Stroop CW.

For the COWAT, we found a medium positive ES of 0·62 ($p=0{\cdot}00$). This indicates that NC participants generated more words during this verbal fluency task than ADHD participants. A medium positive ES of 0·55 ($p=0{\cdot}00$) was also established for attentiveness (d') on the CPT, denoting that the NC group showed a better ability to distinguish important from non-important information on a stimulus level. ADHD participants showed worse inhibition as measured by commission errors on the CPT, reflected in a medium positive ES of 0·64 ($p=0{\cdot}00$) for this variable. For risk taking (β) on the CPT, there was a non-significant ($p=0{\cdot}26$) small negative ES of $-0{\cdot}22$. This indicates that there was no difference in response style (impulsive *versus* cautious) between the ADHD and the NC groups. The ADHD group performed much worse on interference control as measured by the Stroop CW card, as indicated by a large positive ES ($d=0{\cdot}89$, $p=0{\cdot}00$). However, when we controlled the score on the CW card for color naming (the score on card C), there was no difference between the two groups, as indicated by the positive ES of 0·13 ($p=0{\cdot}26$). On TMT-B, a medium positive ES of 0·65 ($p=0{\cdot}00$) could be established, indicating that the NC participants performed better at this set-shifting measure than the ADHD participants. Finally, we found a small positive ES of 0·44 ($p=0{\cdot}01$) for WAIS-DS-BW, implying that the ADHD group has more problems with verbal working memory than the NC group.

Non-EF

The results of the analyses of the non-EF measures are summarized in Table 3. Q values indicated homogeneity for all but two non-EF ES (Stroop W and Stroop C).

For Hit RT on the CPT, there was a non-significant ES of $-0{\cdot}03$ ($p=0{\cdot}79$), which indicates that there were no differences in reaction-time speed for correct responses between the ADHD and the NC groups. The ADHD group showed more variability in reaction times than the NC group, as shown by the medium positive ES of 0·57 ($p=0{\cdot}00$) for Hit RT SE. The medium positive ES of 0·50 ($p=0{\cdot}00$) for omission errors on the CPT points out that the ADHD participants made more of this type of error, suggesting worse vigilance in this group. Both for the Stroop W card and the Stroop C card we observed medium positive ES of 0·60 ($p=0{\cdot}02$) and 0·62 ($p=0{\cdot}01$) respectively. These values imply that the ADHD group had more difficulties than the NC group on both color name reading and color naming. The small positive ES of 0·46 ($p=0{\cdot}00$) for TMT-A denotes that the ADHD group performed poorer than the control group on this measure of serial information processing, visual scanning, and motor speed. A small positive ES of 0·29 ($p=0{\cdot}02$) for WAIS-DS-FW indicates that there is only a small, but significant advantage for the NC group as far as verbal memory span is concerned.

Moderator variables

A major problem in meta-analytic research is the fact that factors other than chance and the cognitive processes under study (EF and non-EF) may influence the difference between groups, especially in the case of heterogeneity in ES. Statistical correction for these factors in a meta-analysis is only sensible with a larger number of studies than was included in the present paper. Therefore, we now discuss several potential moderator variables (see Table 4).

Table 1. *Studies included in the current meta-analysis*

Study	Subjects (% males in sample)	Age (years) Mean (S.D.)	Test and dependent variable
Barkley *et al.* (1996)	ADHD, $n=25$ (64%) NC, $n=23$ (61%)	ADHD, 22·5 (4·0) NC, 22·0 (4·0)	COWAT CPT Hit Reaction Time CPT SE Reaction Time CPT Omissions CPT Attentiveness (d') CPT Commissions CPT Risk Taking (β)
Dinn *et al.* (2001)	ADHD, $n=25$ (36%) NC, $n=11$ (45%)	ADHD, 35·6 (15·9) NC, 35·4 (9·9)	COWAT
Epstein *et al.* (1998)	ADHD, $n=60$ (57%) NC, $n=72$ (58%)	ADHD, 35 (11) NC, 25 (10)	CPT Hit Reaction Time CPT SE Reaction Time CPT Omissions CPT Attentiveness (d') CPT Commissions CPT Risk Taking (β)
Epstein *et al.* (2001)	ADHD, $n=25$ (40%) NC, $n=30$ (50%)	ADHD, 33·6 (–) NC, 33·4 (–)	CPT Hit Reaction Time CPT SE Reaction Time CPT Omissions CPT Attentiveness (d') CPT Commissions CPT Risk Taking (β)
Holdnack *et al.* (1995)	ADHD, $n=25$ (60%) NC, $n=30$ (63%)	ADHD, 30·6 (8·5) NC, 26·7 (6·7)	Trailmaking Test – A
Johnson *et al.* (2001)	ADHD, $n=56$ (71%) NC, $n=38$ (63%)	ADHD, 33·3 (8·42) NC, 40·8 (10·24)	COWAT Stroop Word Stroop Color Stroop Color Word Stroop Interference Trailmaking Test – A Trailmaking Test – B
Lovejoy *et al.* (1999)	ADHD, $n=26$ (50%) NC, $n=26$ (50%)	ADHD and NC range 21–55, median 41	COWAT Trailmaking Test – A Trailmaking Test – B
Murphy (2002)	ADHD, $n=18$ (100%) NC, $n=18$ (100%)	ADHD range 27–58 NC range 25–59	Trailmaking Test – A Trailmaking Test – B
Murphy *et al.* (2001)	ADHD, $n=105$ (75%) NC, $n=64$ (69%)	ADHD, 21·1 (2·7) NC, 21·2 (2·4)	COWAT CPT Hit Reaction Time CPT SE Reaction Time CPT Omissions CPT Attentiveness (d') CPT Commissions CPT Risk Taking (β) WAIS-DS-FW WAIS-DS-BW
Rapport *et al.* (2001)	ADHD, $n=35$ (69%) NC, $n=32$ (59%)	ADHD, 32·9 (10·8) NC, 33·2 (13·2)	COWAT Trailmaking Test – A Trailmaking Test – B
Riordan *et al.* (1999)	ADHD, $n=21$ (81%) NC, $n=15$ (47%)	ADHD, 31·8 (11·8) NC, 36·5 (10·8)	COWAT Stroop Word Stroop Color Stroop Color Word Stroop Interference WAIS-DS-FW WAIS-DS-BW Trailmaking Test – A Trailmaking Test – B
Taylor & Miller (1997)	ADHD, $n=211$ (–) NC, $n=28$ (–)	—	Stroop Word Stroop Color Stroop Color Word Stroop Interference Trailmaking Test – A Trailmaking Test – B
Walker *et al.* (2000)	ADHD, $n=30$ (83%) NC, $n=30$ (67%)	ADHD, 25·8 (8·7) NC, 25·8 (6·8)	COWAT CPT Hit Reaction Time CPT SE Reaction Time CPT Omissions CPT Commissions Stroop Word Stroop Color Stroop Color Word Stroop Interference WAIS-DS-FW WAIS-DS-BW Trailmaking Test – A Trailmaking Test – B

Dashes indicate that information was not provided in original paper. ADHD, Attention deficit hyperactivity disorder; COWAT, Controlled Oral Word Association Test; CPT, Continuous Performance Test; NC, Normal Control; S.E., standard error; WAIS-DS-FW, WAIS Digit Span Forwards; WAIS-DS-BW, WAIS Digit Span Backwards.

Table 2. *Combined random effect sizes and statistical outcomes for EF measures*

Measure	ADHD ($n=$)	NC ($n=$)	Cohen's d	t value	p value	Lower limit	Upper limit	Q value
COWAT	323	239	0·62	3·74	0·00	0·30	0·94	22·01*
CPT attentiveness (d')	215	189	0·55	5·35	0·00	0·35	0·75	0·71
CPT commissions	245	219	0·64	5·26	0·00	0·40	0·88	5·86
CPT risk taking (β)	215	189	−0·22	−1·13	0·26	−0·61	0·16	9·90*
Stroop CW	318	111	0·89	3·19	0·00	0·34	1·44	13·94*
Stroop Interference	318	111	0·13	1·14	0·26	−0·10	0·37	1·50
TMT-B	397	187	0·65	6·67	0·00	0·46	0·85	3·40
WAIS-DS-BW	156	109	0·44	2·57	0·01	0·10	0·78	3·03

ADHD, Attention deficit hyperactivity disorder; COWAT, Controlled Oral Word Association Test; CPT, Continuous Performance Test; NC, Normal Control; Stroop CW, Stroop Color Word Card; TMT-B, Trailmaking Test – Part B; WAIS-DS-BW, WAIS Digit Span Backwards.

* Indicates heterogeneity of effect sizes ($p<0·05$).

Table 3. *Combined random effect sizes and statistical outcomes for non-EF measures*

Measure	ADHD ($n=$)	NC ($n=$)	Cohen's d	t value	p value	Lower limit	Upper limit	Q value
CPT Hit RT	245	219	−0·03	−0·26	0·79	−0·22	0·17	4·47
CPT SE RT	245	219	0·57	4·14	0·00	0·30	0·83	7·31
CPT Omissions	245	219	0·50	5·00	0·00	0·31	0·70	4·32
Stroop W	318	111	0·60	2·43	0·02	0·11	1·08	11·32*
Stroop C	318	111	0·62	2·80	0·01	0·18	1·06	9·25*
TMT-A	422	217	0·46	4·96	0·00	0·28	0·65	7·27
WAIS-DS-FW	156	109	0·29	2·32	0·02	0·04	0·54	1·13

ADHD, Attention deficit hyperactivity disorder; CPT, Continuous Performance Test; NC, Normal Control; RT, Reaction Time; S.E., standard error; Stroop C, Stroop Color Card; Stroop W, Stroop Word Card; TMT-A, Trailmaking Test – Part A; WAIS-DS-FW, WAIS Digit Span Forwards.

* Indicates heterogeneity of effect sizes ($p<0·05$).

First of all, the studies differed with respect to the diagnostic procedures for ADHD. One of the problems in diagnosing adult ADHD is that symptoms have to have started before the age of 7 years. This means retrospectively establishing those symptoms, which raises questions of reliability and validity of the diagnosis. Another concern is the reliability of patient self-reports about their symptoms (Barkley *et al.* 2002). Therefore, to reduce the chance of both false positives and false negatives, it is best if more than one informant is consulted (e.g. the patient, a parent, a spouse) and if more than one type of measurement is used (e.g. self-report questionnaires, clinical interviews, structured interviews; Weiss & Murray, 2003). Next to heterogeneity between samples, ADHD in itself is a heterogeneous diagnosis with many different symptoms leading to several different subtypes, which also complicates comparing studies.

Another confounder can be found in the fact that ~75% of adults with ADHD suffer from other psychiatric disorders as well (Biederman *et al.* 1993). Many of these disorders may also be attended with cognitive disabilities, so that it is hard to conclude if established difficulties in cognitive areas are related to the ADHD or to the co-existing disorder. Ideally, participants should be tested for co-existing disorders and there should be some form of statistical correction for this co-morbidity.

Thirdly, men and women differ in their cognitive abilities (Kimura, 1996), so if the composition of the ADHD group and the NC group differs with respect to sex, this may influence the results. Also, it may not be possible to compare studies when some have included only men, and others have also tested women.

A fourth possible moderator variable is the intelligence level of participants. There is continuing debate in the current literature as to whether EF data should be corrected for overall IQ level (Denckla, 1996). Especially in children with ADHD, many researchers have noted a

Table 4. *Potential moderator variables*

Study	ADHD diagnosis	Subtypes	Co-morbid disorders	IQ	Medication
Barkley *et al.* (1996)	1 informant >1 measurement	100% combined	Investigated	No difference	Testing after washout
Dinn *et al.* (2001)	1 informant >1 measurement	52% combined 16% H/I 32% I	Investigated	—	Half of sample on medication, differences with unmedicated group only for one test
Epstein *et al.* (1998)	1 informant >1 measurement	23% combined 12% H/I 65% I	—	—	—
Epstein *et al.* (2001)	1 informant >1 measurement	40% combined 4% H/I 56% I	Investigated	—	Unmedicated
Holdnack *et al.* (1995)	1 informant >1 measurement	—	—	Difference statistically controlled for	Unmedicated
Johnson *et al.* (2001)	1 informant >1 measurement	—	Investigated	Results with and without controlling for IQ	Testing after washout
Lovejoy *et al.* (1999)	1 informant >1 measurement	—	Investigated	No difference	Testing after washout
Murphy (2002)	1 informant 1 measurement	100% combined	Investigated	No difference	—
Murphy *et al.* (2001)	>1 informant >1 measurement	55% combined 2% H/I 34% I 9% NOS	Investigated and statistically controlled for	Results with and without controlling for IQ	Testing after washout
Rapport *et al.* (2001)	1 informant >1 measurement	—	Investigated	No difference	Unmedicated
Riordan *et al.* (1999)	1 informant >1 measurement	—	Investigated	Difference statistically controlled for	Unmedicated
Taylor & Miller (1997)	>1 informant >1 measurement	57% combined 3% H/I 38% I 2% NOS	Investigated	—	—
Walker *et al.* (2000)	>1 informant >1 measurement	—	Investigated	No difference	Unmedicated

Dashes indicate that information was not provided in original paper. ADHD, Attention deficit hyperactivity disorder; H/I, Hyperactive/Impulsive subtype; I, Inattentive subtype; NC, Normal Control; NOS, not otherwise specified.

correlation between EFs and IQ (e.g. Ardila *et al.* 2000), indicating at least a relation between the two. Other researchers (e.g. Nigg, 2001) have argued that controlling for IQ might remove some of the variance that is related to ADHD. Ideally, researchers should, therefore, report their EF results with and without controlling for overall IQ performance (Barkley, 1997*a*). This was done in only two of the 13 studies used for this meta-analysis (Murphy *et al.* 2001; Johnson *et al.* 2001).

Next, the medication of choice for ADHD (methylphenidate) is known to have an effect on several cognitive abilities, both in children (e.g. Tannock *et al.* 1995) and in adults with the disorder (e.g. Kuperman *et al.* 2001). However, in three studies included in this review, it was not even mentioned whether ADHD participants were taking medication or not (Taylor & Miller, 1997; Epstein *et al.* 1998; Murphy, 2002).

Finally, one would preferably want to compare the ADHD group with a group of NCs that, in line with the argumentation for other moderating variables, shows no signs of psychopathology, does not take any kind of psychotropic medication, and is of similar gender, age and IQ as the ADHD group. The NC groups in the studies included here vary largely. In some studies, the criteria for the NC groups remain vague (Epstein *et al.* 1998; Dinn *et al.* 2001). Most researchers clearly state that NC participants were not allowed to score above a certain cut-off score on some measure for ADHD, although childhood ADHD was not always an exclusion criterion (Johnson *et al.* 2001). Neurological conditions or events and other psychiatric diagnoses were usually reason for exclusion, although studies varied in the ways of establishing these other diagnoses (by clinical interview, structured interview, self-report or questionnaire). In the study by Taylor & Miller (1997), the 'No Diagnoses' group consisted of people who were self-referred for evaluation of ADHD, but who did not meet DSM-IV criteria for ADHD. It could of course be questioned as to how compatible this latter group was to other NC groups, and even if this group would not be more like the ADHD group than like a NC group.

DISCUSSION

We conducted the present study to establish a quantitative account of the difference in EF between adults with ADHD and NCs. We included non-EF dependent variables from the EF tasks, in order to determine whether deficits are specific to EF or not. As far as we know, this study is one of the first quantitative reviews of this topic, and based on the average number of subjects for each analysis, the analyses had enough power to be able to draw some firm conclusions.

Our results in the EF domain are in agreement with the child literature on ADHD, where differences between children with ADHD and NCs in the areas of verbal fluency, inhibition, and set shifting have been reported consistently (Sergeant *et al.* 2002). In their qualitative review, Woods *et al.* (2002) concluded that Stroop tasks, verbal letter fluency, auditory verbal list learning, and continuous performance tests discriminate best between adult ADHD and NC samples. Our data provide no answers with respect to auditory verbal list learning, since insufficient data were available for these analyses. With respect to Stroop tasks, our data demonstrated that people with ADHD show worse performance than NCs on all three cards of the Stroop, not just on the interference (CW) card. When controlling for performance on the Color card, the ES for the Color Word card was no longer significant. Therefore, we cannot conclude that adults with ADHD show poor selective visual attention and/or prepotent response inhibition, as Woods *et al.* (2002) suggested. Future research including the Stroop Color Word Test should correct for performance on at least the Color card when reporting interference results for this test. With respect to the TMT, Woods *et al.* (2002) concluded that many studies have shown differences on part A, and not so much on part B. According to these authors, this may be related to the initial novelty of the task. Our quantitative analyses are partly in agreement with this point, since we found a small ES for TMT-A. However, we found a larger ES for TMT-B, indicating more robust differences on this part of the test. Based on our data, it might be concluded that there is a set-shifting problem in adult ADHD, and not just a problem with novelty. To be able to draw firmer conclusions in this area, it will be necessary to correct performance on part B for performance on part A, as was done with the Stroop Color Word Test. However, the data to perform these analyses were not available. It would make sense for future studies to correct performance on part B for performance on part A, by looking at difference scores. The same advice holds for WAIS Digit Span, where performance on DS Backwards should be corrected for performance on DS Forwards, before conclusions with respect to working memory can be drawn, based on performance on this test. With respect to verbal fluency tests, our data are in agreement with the conclusion by Woods *et al.* (2002), however, we do not feel that these tests 'demonstrate great promise in discriminating adults with ADHD from comparison groups' (p. 22), since other psychiatric groups have been shown to perform poorly on this type of measure and it thus lacks specificity (e.g. Harvey *et al.* 1997).

In the non-EF domain, variability in reaction times has been noted before in relation to ADHD, both in adults (Tinius, 2003) and in children (Scheres *et al.* 2001). Inconsistency has also been noted in other areas of performance in ADHD, such as motor timing (Rubia *et al.* 1999). This 'consistent inconsistency' may well be related to the recent suggestion of an endophenotype (intermediate construct between genes and behavior) in ADHD related to variability in performance (Castellanos & Tannock, 2002). Although this endophenotype is connected to inter-individual variability, rather than variability between subjects, it is noteworthy that the measures with large ES (COWAT and Stroop) are also the measures with significant *Q* values. This indicates that also within ADHD as a group, performance may not be consistent. Poorer performance on the other tasks (Stroop, CPT Omissions, TMT-A, WAIS-DS-FW) has been noted before in children and adults with ADHD. Many of these variables seem to point towards general slowing on more cognitive responses (like reading, color naming, and visual search), even though motor response as measured by CPT Hit RT is not slower. This general cognitive slowing, as opposed to motor slowing, is in line with earlier research (e.g. Aldenkamp *et al.* 2000). Verbal memory deficits (WAIS-DS-FW) have also been noted in ADHD before (Quinlan & Brown, 2003).

In light of the current emphasis on EF in ADHD research, we feel that the most striking outcome of this review is the similarity in ES between the EF domain and the non-EF domain. Simply averaging the ES for both domains yielded a mean ES of 0·40 for the EF variables (we excluded Stroop CW in this calculation) and a mean ES of 0·43 for the non-EF domain. The total sample size of the groups compared was large enough to be able to conclude that these figures do not suggest a specific deficit in the EF realm for adults with ADHD. Rather, they suggest that in comparison with NC adults, adults with ADHD show disabilities in various areas of cognitive functioning, including EF. This conclusion needs to be strengthened by analyzing other tests specifically designed to measure non-EF functions, rather than including non-EF dependent variables from EF tests. Nevertheless, the lack of difference between EF and non-EF ES calls into question models of ADHD that depend heavily upon EF for their explanatory power, such as the model by Barkley (1997*b*).

Another striking result from our study, which supports the last statement, is the fact that we found only one large ES, for interference control as measured by the Stroop CW card. However, this ES was no longer significant when we controlled for another function necessary to perform appropriately on this test (color naming). So in fact we only detected medium ES. Cohen (1988) noted that values of *f* as large as 0·50 (corresponding with *d* values of 1·00) are not common in behavioral science, but an area that has received so much attention in research during the past decade might be expected to yield larger ES. Moreover, the largest ES were also the ones that were accompanied by significant *Q* values, indicating heterogeneity in results. This points to the fact that although EF problems are part of ADHD in adults, they are not so in every study and every sample. Again the question arises: Should we continue the quest for EF difficulties in ADHD?

The issue of specificity in EF research also underlines this last question. Sergeant *et al.* (2002) concluded that the EF problems are not specific for ADHD in children, since other psychopathological groups also showed problems with these abilities. Unfortunately, there are only very few studies in adult ADHD that have included clinical comparison groups. The few studies available suggest lack of specificity in adult ADHD as well (Taylor & Miller, 1997; Walker *et al.* 2000; Epstein *et al.* 2001). It is well known that many other psychiatric disorders are accompanied by EF deficits, such as schizophrenia (Velligan & Bow-Thomas, 1999), and depression (Ottowitz *et al.* 2002). Future research urgently needs to employ multiple clinical groups. Especially disorders that either have symptoms in common with ADHD (like depression or mania) or that share involvement of neurotransmitters or frontal areas with ADHD (e.g. schizophrenia) should be compared with ADHD.

EFs have played a major role in many theoretical accounts of ADHD. Although these accounts have not been specifically proposed for ADHD in adults, we would expect them to be applicable to the adult version of the disorder. In line with Pennington & Ozonoff (1996), we would expect primarily deficits in the realm of behavioral inhibition and working memory, whereas according to the theory of Barkley (1997*b*), a core deficit in inhibition would lead to problems in all other areas of EF. Our data support neither view. Various researchers have made other suggestions with regard to a theoretical explanation of ADHD. As mentioned before, some have suggested general slowing as an explanation. This suggestion seems to be backed up by our data. Other researchers have suggested motivational issues and delay aversion, either singly or in combination with inhibition (Sonuga-Barke, 2002) and the role of reward (Douglas, 1999). Unfortunately, no studies have been performed in this area with an adult ADHD population. This also holds for the role of energetics, which has been suggested by Sergeant & van der Meere (1990). More recently, Castellanos & Tannock (2002) argued that one of the key characteristics of ADHD might be the temporal and contextual variability in performance, related to cerebellar dysfunction. Our results support variability in responses (medium ES for CPT standard error of reaction time).

We do not believe that our similar results in the EF and non-EF domains indicate that we should discard the possible EF explanation for ADHD altogether, but it seems high time for some changes in the field. For one thing, it seems, now more than ever, necessary to develop reliable and valid measures of EF. As long as we do not have improved EF measures at our disposal, researchers could improve their efforts by using tests that include different levels of difficulty (like the Tower of London), or that manipulate different functions at the same time. Another way of improving research in this area, is by including control tasks for skills that are not related to EF *per se*, but that are necessary to perform an EF test anyway. It would also be an improvement to use tasks that are based on theoretical accounts of specific cognitive processes, rather than tasks that have been defined as EF task based on lesion studies. Examples of such tasks are the Stop Signal Test (Logan *et al.* 1984), and the Self Ordered Pointing Test (Petrides & Milner, 1982).

To conclude this discussion, we would like to point out some limitations of our study. The first one can be found in the potential moderator variables, of which we provided a detailed overview in the Methods section. Without statistical controls for the effects of the variables, their impact is not quantified and their possible influence should be kept in mind while interpreting our results. Future studies of adult ADHD should aim for careful diminution of methodological differences by taking these issues into account. The second limitation can be found in another well-known problem in meta-analysis: the 'file drawer problem'. This refers to the fact that studies without significant group differences tend to remain in file drawers rather than to get published. This may of course greatly limit the conclusions that can be drawn. Finally, our inclusion criteria of at least four studies with an adult ADHD population and a total number of participants exceeding 50 led to exclusion of some interesting and important papers in the field, which we hope will stimulate further research (e.g. McLean *et al.* 2004).

In sum, in this meta-analytic review we showed differences between adult ADHD and NC in both areas of EF and areas of non-EF. This result raises doubts about the current emphasis on EF research in ADHD. We feel that we should not view the EF research venue as a dead end yet, but that the field is in need of some important methodological changes before we can decide in favor of or against the EF hypothesis of ADHD.

ACKNOWLEDGEMENTS

We thank Dr Russell Barkley, Dr Wayne Dinn, Dr Jeffrey Epstein, Dr Laura Flashman, Dr Diane Johnson, Dr Lisa Rapport, Dr Cindy Taylor and Dr Alexandra Walker for providing additional data or information for this meta-analysis.

DECLARATION OF INTEREST

None.

REFERENCES

Aldenkamp, A., van Bronswijk, K., Braken, M., Diepman, L. A., Verwey, L. E. & van den Wittenboer, G. (2000). A clinical comparative study evaluating the effect of epilepsy versus ADHD on timed cognitive tasks in children. *Child Neuropsychology* **6**, 209–217.

Ardila, A., Pineda, D. & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology* **15**, 31–36.

Barkley, R. A. (1997*a*). *ADHD and the nature of self control.* Guilford Press: New York.

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

Barkley, R. A., Fischer, M., Smallish, L. & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology* **111**, 279–289.

Barkley, R. A., Murphy, K. & Kwasnik, D. (1996). Psychological adjustments and adaptive impairments in young adults with ADHD. *Journal of Attention Disorders* **1**, 41–54.

Biederman, J., Faraone, S. V., Spencer, T., Wilens, T., Norman, D., Lapey, K. A., Mick, E., Lehman, B. K. & Doyle, A. (1993). Patterns of psychiatric comorbidity, cognition, and psychosocial functioning in adults with attention deficit hyperactivity disorder. *American Journal of Psychiatry* **150**, 1792–1798.

Borenstein, M. & Rothstein, H. (1999). *Comprehensive Meta Analysis: A Computer Program for Research Synthesis*. Biostat: Englewood, NJ.

Castellanos, F. X. & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews. Neuroscience* **3**, 617–628.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd edn). Lawrence Erlbaum: Hillsdale, NJ.

Conners, C. K. (1995). Conners' Continuous Performance Test computer program. Multi Health Systems Inc.: Ontario, Canada.

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**, 114–119.

Dinn, W. M., Robbins, N. C. & Harris, C. L. (2001). Adult attention-deficit/hyperactivity disorder: neuropsychological correlates and clinical presentation. *Brain and Cognition* **46**, 114–121.

Douglas, V. I. (1999). Cognitive control processes in attention-deficit/hyperactivity disorder. In *Handbook of Disruptive Behavior Disorders* (ed. H. C. Quay and A. E. Hogan), pp. 105–138. Plenum Press: New York.

Epstein, J. N., Conners, C. K., Sitarenios, G. & Erhardt, D. (1998). Continuous Performance Test results of adults with Attention Deficit Hyperacitivity Disorder. *The Clinical Neuropsychologist* **12**, 155–168.

Epstein, J. N., Johnson, D. E., Varia, I. M. & Conners, C. K. (2001). Neuropsychological assessment of response inhibition in adults with ADHD. *Journal of Clinical and Experimental Neuropsychology* **23**, 362–371.

Hammes, J. G. W. (1971). *De Stroop Kleur-Woord Test. Handleiding*. Swets & Zeitlinger: Lisse, The Netherlands.

Harvey, P. D., Lombardi, J., Leibman, M., Parrella, M., White, L., Powchik, P., Mohs, R. C. & Davidson, M. (1997). Verbal fluency deficits in geriatric and nongeriatric chronic schizophrenic patients. *Journal of Neuropsychiatry and Clinical Neurosciences* **9**, 584–590.

Hedges, L. V. & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Academic Press: New York.

Holdnack, J. A., Moberg, P. J., Arnold, S. E., Gur, R. C. & Gur, R. E. (1995). Speed of processing and verbal learning deficits in adults diagnosed with attention deficit disorder. *Neuropsychiatry, Neuropsychology and Behavioral Neurology* **8**, 282–292.

Johnson, D. E., Epstein, J. N., Waid, L. R., Latham, P. K., Voronin, K. E. & Anton, R. F. (2001). Neuropsychological performance deficits in adults with attention deficit/hyperactivity disorder. *Archives of Clinical Neuropsychology* **16**, 587–604.

Kimura, D. (1996). Sex, sexual orientation and sex hormones influence human cognitive function. *Current Opinion in Neurobiology* **6**, 259–263.

Kulik, J. A. & Kulik, C. L. C. (1989). Meta-analysis in education. *International Journal of Educational Research* **13**, 227–340.

Kuperman, S., Perry, P. J., Gaffney, G. R., Lund, B. C., Bever-Stille, K. A., Arndt, S., Holman, T. L., Moser, D. J. & Paulsen, J. S. (2001). Bupropion SR versus methylphenidate versus placebo for attention deficit hyperactivity disorder in adults. *Annals of Clinical Psychiatry* **13**, 129–134.

Logan, G. D., Cowan, W. B. & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *Journal of Experimental Psychology. Human Perception and Performance* **10**, 276–291.

Lovejoy, D. W., Ball, J. D., Keats, M., Stutts, M. L., Spain, E. H., Janda, L. & Janusz, J. (1999). Neuropsychological performance of adults with attention deficit hyperactivity disorder (ADHD): diagnostic classification estimates for measures of frontal lobe/executive functioning. *Journal of the International Neuropsychological Society* **5**, 222–233.

McLean, A., Dowson, J., Toone, B., Young, S., Bazanis, E., Robbins, T. W. & Sahakian, B. J. (2004). Characteristic neurocognitive profile associated with adult attention-deficit/hyperactivity disorder. *Psychological Medicine* **34**, 681–692.

Murphy, K. R., Barkley, R. A. & Bush, T. (2001). Executive functioning and olfactory identification in young adults with attention deficit-hyperactivity disorder. *Neuropsychology,* **15(2)**, 211–220.

Murphy, P. (2002). Cognitive functioning in adults with Attention-Deficit/Hyperactivity Disorder. *Journal of Attention Disorders* **5**, 203–209.

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

Ottowitz, W. E., Dougherty, D. D. & Savage, C. R. (2002). The neural network basis for abnormalities of attention and executive function in major depressive disorder: implications for application of the medical disease model to psychiatric disorders. *Harvard Review of Psychiatry* **10**, 86–99.

Pennington, B. F. & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry* **37**, 51–87.

Petrides, M. & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* **20**, 249–262.

Quinlan, D. M. & Brown, T. E. (2003). Assessment of short-term verbal memory impairments in adolescents and adults with ADHD. *Journal of Attention Disorders* **6**, 143–152.

Rapport, L. J., Van Voorhis, A., Tzelepis, A. & Friedman, S. R. (2001). Executive functioning in adult attention-deficit hyperactivity disorder. *The Clinical Neuropsychologist* **15**, 479–491.

Reitan, R. M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery*. Neuropsychology Press: Tuscon.

Riordan, H. J., Flashman, L. A., Saykin, A. J., Frutiger, S. A., Carroll, K. E. & Huey, L. (1999). Neuropsychological correlates of methylphenidate treatment in adult ADHD with and without depression. *Archives of Clinical Neuropsychology* **14**, 217–233.

Rubia, K., Taylor, A., Taylor, E. & Sergeant, J. A. (1999). Synchronization, anticipation, and consistency in motor timing of children with dimensionally defined attention deficit hyperactivity disorder. *Perceptual and Motor Skills* **89**, 1237–1258.

Scheres, A., Oosterlaan, J. & Sergeant, J. A. (2001). Response execution and inhibition in children with AD/HD and other disruptive disorders: the role of behavioural activation. *Journal of Child Psychology and Psychiatry* **42**, 347–357.

Sergeant, J. A., Geurts, H. M. & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for Attention-Deficit/Hyperactivity Disorder? *Behavioural Brain Research* **130**, 3–28.

Sergeant, J. A. & van der Meere, M. J. (1990). Additive factor method applied to psychopathology with special reference to childhood hyperactivity. *Acta Psychologica* **74**, 277–295.

Sonuga-Barke, E. J. (2002). Psychological heterogeneity in AD/HD – a dual pathway model of behaviour and cognition. *Behavioural Brain Research* **130**, 29–36.

Spreen, O. & Benton, A. L. (1977). *The Neurosensory Center Comprehensive Examination for Aphasia*. Victoria, Canada: University of Victoria, Neuropsychology Laboratory.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* **18**, 643–662.

Tannock, R., Schachar, R. & Logan, G. (1995). Methylphenidate and cognitive flexibility: dissociated dose effects in hyperactive children. *Journal of Abnormal Child Psychology* **23**, 235–266.

Taylor, C. J. & Miller, D. C. (1997). Neuropsychological assessment of attention in adults. *Journal of Attention Disorders* **2**, 77–88.

Tinius, T. P. (2003). The Integrated Visual and Auditory Continuous Performance Test as a neuropsychological measure. *Archives of Clinical Neuropsychology* **18**, 439–454.

Velligan, D. I. & Bow-Thomas, C. C. (1999). Executive function in schizophrenia. *Seminars in Clinical Neuropsychiatry* **4**, 24–33.

Walker, A. J., Shores, E. A., Trollor, J. N., Lee, T. & Sachdev, P. S. (2000). Neuropsychological functioning of adults with attention deficit hyperactivity disorder. *Journal of Clinical and Experimental Neuropsychology* **22**, 115–124.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised*. The Psychological Corporation: New York.

Weiss, M. & Murray, C. (2003). Assessment and management of attention-deficit hyperactivity disorder in adults. *Journal of the Canadian Medical Association* **168**, 715–722.

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

Woods, S. P., Lovejoy, D. W. & Ball, J. D. (2002). Neuropsychological characteristics of adults with ADHD: a comprehensive review of initial studies. *The Clinical Neuropsychologist* **16**, 12–34.