Booth, Josephine N. and Boyle, James and Kelly, Steve (2010) Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: findings from a meta-analysis. British Journal of Developmental Psychology, 28 (1). pp. 133-176. ISSN 0261-510X

http://strathprints.strath.ac.uk/25714/

Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis

Josephine N. Booth*, James M. E. Boyle and Steve W. Kelly
Department of Psychology, University of Strathclyde, Glasgow, UK

Research studies have implicated executive functions in reading difficulties (RD). But while some studies have found children with RD to be impaired on tasks of executive function other studies report unimpaired performance. A meta-analysis was carried out to determine whether these discrepant findings can be accounted for by differences in the tasks of executive function that are utilized. A total of 48 studies comparing the performance on tasks of executive function of children with RD with their typically developing peers were included in the meta-analysis, yielding 180 effect sizes. An overall effect size of 0.57 (*SE* .03) was obtained, indicating that children with RD have impairments on tasks of executive function. However, effect sizes varied considerably suggesting that the impairment is not uniform. Moderator analysis revealed that task modality and IQ-achievement discrepancy definitions of RD influenced the magnitude of effect; however, the age and gender of participants and the nature of the RD did not have an influence. While the children's RD were associated with executive function impairments, variation in effect size is a product of the assessment task employed, underlying task demands, and definitional criteria.

It is estimated that between 10 and 15% of schoolchildren have difficulties with reading (Velluntino & Fletcher, 2007). While theoretical accounts of reading difficulties (RD) posit the primary deficit to be in the phonological system (Hulme & Snowling, 2009; Velluntino, Fletcher, Snowling, & Scanlon, 2004), deficiencies in the executive system have also been identified (Pickering & Gathercole, 2004; Swanson, 2006) and it has been suggested that these problems could be 'above and beyond their deficits in the phonological system' (Swanson, 2006, p. 58). Swanson (2006) highlights several executive areas where children with RD have difficulties, including maintaining relevant information in working memory, inhibition of irrelevant

information, and accessing material in long-term memory. Furthermore, theoretical accounts of difficulties with reading comprehension implicate working memory skills, comprehension monitoring, and inference making (Hulme & Snowling, 2009; Nation, 2007; Vukovic & Siegel, 2006), and the ability to update information and inhibit distractors has been implicated in research (e.g. Palladino, Cornoldi, De Beni, & Pazzaglia, 2001). These 'higher order' cognitive processes (Rayner & Pollatsek, 1989) all fall under the rubric of executive functions and may have implications for differential responsivity to reading intervention (Sesma, Mahone, Levine, Eason, & Cutting, 2009).

Executive functions and reading

Executive functions are defined as the underlying processes involved in cognitive functioning (Booth, Boyle, & Kelly, in press). There is a growing body of research which has implicated executive functions in many areas of language learning; for example, the learning of new vocabulary (Dempster & Cooney, 1982), language abilities (Ellis Weismer, Evans, & Hesketh, 1999), literacy (Gathercole & Pickering, 2000), sentence reading (Gernsbacher, 1993), and language and reading comprehension (Booth *et al.*, in press; Daneman & Merikle, 1996; Dempster & Corkill, 1999; Gathercole & Pickering, 2000; Palladino *et al.*, 2001; Sesma *et al.*, 2009).

Several studies have found that children with RD are impaired on tasks of executive functions. For example, Everatt, Warner, Miles, and Thomson (1997) found that children with RD were impaired on the Stroop task compared to typically developing controls (TDC). Martinussen and Tannock (2006) found impaired performance on tasks of verbal and visuospatial working memory for children with RD. Furthermore, Miller-Shaul (2005) gave children with RD a battery of tasks assessing different aspects of executive function and found statistically lower performance of children with RD across all areas. However, some studies have found that the performance of children with RD is comparable to that of TDC. For example, McGee, Brodeur, Symons, Andrade, and Fahie (2004) investigated differences between children with RD and clinical control children as part of a larger study looking at the dissociation between attention-deficit hyperactivity disorder (ADHD) and RD. They found there was no statistical difference in performance between the RD group and the control group on tasks of working memory and also on the Conners' Continuous Performance Task (Conners, 1995), which gives an overall index which is indicative of attention problems. Furthermore, Swanson, Saez, and Gerber (2004), conducted a large-scale study assessing the predictive value of phonological and executive processes on later reading performance. As part of this study, children who scored at least one standard deviation below the mean on a task of word reading were compared to those who scored above this cut-off score on a battery of tasks of working memory and also random generation tasks which are designed to tap inhibitory skills. No significant difference on task performance was found between these two groups.

A possible source of confound, which may underpin the discrepancies between studies which have found significant differences between RD and control groups and studies which have not, is the variety of tasks of executive function which have been employed. The purpose of the present meta-analysis, therefore, is to assess whether the discrepancies found in the literature are, at least partly, a consequence of the assessment tasks employed.

Variety of constructs

The term executive function is generally seen to encompass a wide array of processes; for example, Baron (2004) lists 21 different subdomains of executive functions, including processes such as problem solving and attentional control. While theoretical accounts have placed emphasis on different processes within the umbrella of executive functioning it has been suggested that inhibition and working memory may be integral (Barkley, 1997; Denckla, 1996; Diamond, Barnett, Thomas, & Munro, 2007; Pennington, Bennetto, McAleer, & Roberts, 1996; Roberts & Pennington, 1996).

Factor analytic studies have found that tasks of executive function load on to several distinct factors, namely inhibition, working memory, and in some cases shifting (St Clair-Thompson & Gathercole, 2006; Willcutt *et al.*, 2001; Willcutt, Pennington, Olson, Chhabildas, & Hulslander, 2005). However, in addition to finding several distinct factors, research by Miyake, Friedman, Emerson, Witzki, and Howerter (2000) found that executive functions are also highly related. They focused on three executive functions of shifting, updating of working memory representations, and inhibition. They reported evidence for the independent contribution of each of these constructs to performance on tasks commonly used to tap executive function, thereby showing that these three functions are separable constructs. However, shifting, updating, and inhibition were also found to be significantly correlated with each other and Miyake concluded that this demonstrated both the 'unity and diversity' of executive functions and suggested that underlying inhibitory processes may be involved in all three of the executive functions investigated. Further support for this unity and diversity was found by Lehto, Juujärvi, Kooistra, and Pulkkinen (2003) who reported a similar pattern of results with a sample of children. Therefore while evidence suggests that executive functions are all distinct constructs, it seems that they are not entirely unrelated.

Measurement difficulty

A range of measurement tasks have been used in the assessment of executive function and difficulty arises from the fact that it is often unclear exactly which areas of executive function are being measured by which task, and indeed how much each task assesses multiple constructs. This is generally termed 'task impurity' (Miyake *et al.*, 2000; Rabbitt, 1997; van der Sluis, de Jong, & van der Leij, 2004). That is, that many tasks commonly used to tap executive functions actually involve more than one executive process. It is possible that this task impurity issue could make isolating areas of executive difficulty problematic, which could influence our understanding of the role of executive function in RD.

In order to investigate the issue of task impurity, Morris (1996) conducted a survey of measures of executive function. Six prominent journals were screened for measures of 'executive function' (defined by Morris as involving problem-solving skills), attention, and memory used with children of school age. It was found that in excess of 20 measures of 'executive function' were described, 15 measures of memory, and more than 25 measures of attention. One of the most conspicuous findings of this survey was that many of the measures of 'executive function' were used by other researchers as measures of attention. That is, that there was little consensus about the underlying processes which are measured by tasks.

Further investigation of the tasks identified led Morris (1996, p. 13) to conclude that 'a majority of the tests described as measuring a single construct were actually multidimensional in nature'. This finding is supported by Ozonoff (1997) who gave

examples of several of the most widely used measures of executive function (i.e. the Wisconsin Card Sorting Test (WCST), the Tower of Hanoi, the Matching Family Figures Test, the Trial Making Test, and the Stroop task) as measures which involve more than one executive function. Indeed, Rabbitt (1997, p. 13) suggests that 'executive tasks are, necessarily, very complex, and that attempts to fit them into linguistic categories borrowed from everyday discourse such as "inhibition", or "planning", or "monitoring", are necessarily Procrustean'. That is, it is inappropriate to attempt to define these tasks as assessing one singular construct. Understandably then this poses problems for the interpretation of poor performance on these tasks and complicates isolating areas of possible executive dysfunction. This is further confounded by the lack of consensus regarding which areas of executive function are measured by each task and makes the interpretation of poor performance on these tasks extremely difficult.

Comorbidity is a further complicating factor in the study of executive functions in children with RD. As RD are known to co-occur with many other developmental disorders, such as ADHD (Tannock, 1998), the identification of areas of difficulty specific to RD is not always straightforward. Executive function difficulties are thought to be one of the main characteristics of ADHD (Hulme & Snowling, 2009). Research by Willcutt *et al.* (2001, 2005) examined the executive function profile of groups of children with ADHD only, RD only, comorbid ADHD and RD, and neither disorder. While the RD group and the ADHD group had differing areas of executive function difficulty, both groups showed more symptoms of the other disorder than the control group. That is, while the RD group did not meet clinical cut-offs on assessment of ADHD, they did show more symptoms than the control group and vice versa with the ADHD group showing symptoms of RD. This therefore indicates that groups of children with RD may display subclinical presentations of other disorders. As executive function difficulties are thought to occur in both of these disorders, it is difficult to attribute executive impairments solely to RD thus complicating further the interpretation of poor executive function task performance.

The present study: Candidate moderators

The primary objective of the present meta-analysis is to evaluate the variety of tasks of executive function which have been identified in the literature regarding children with RD. As such several candidate moderator variables will be explored. Firstly, the impact of IQ-discrepant criteria in defining RD will be examined as a moderator, followed by the nature of the RD, that is, whether it centres around word reading or reading comprehension. Further to this, both age and gender will be examined as candidate moderators and then finally the modality of the measurement task. Each of these moderators will now be explored in turn.

The impact of IQ

Stuebing *et al.* (2002) reviewed the applicability of using an IQ-achievement discrepancy definition of RD in several key areas, including executive function, where RD is defined as a significant discrepancy between IQ and reading attainment. The literature which had directly compared IQ-discrepant readers with IQ-consistent (or non-discrepant) readers was examined and effect sizes relating to behaviour, achievement, and cognitive ability measures were calculated. A medium effect size (0.41) was found for measures of executive function in favour of children whose reading

problems had been defined using IQ-discrepancy criteria. However, the authors suggested that the effect sizes were merely a product of these definitional criteria, i.e. the relationship with IQ, and not due to any real differences in executive performance between IQ-discrepant and IQ-consistent readers. They suggested that the use of a discrepancy definition did not add to our understanding of RD and concluded the review by arguing against the use of an IQ-discrepancy definition. Despite this finding, IQ-discrepancy definitions are still employed in research which investigates executive functioning and RD (e.g. Altemeier, Abbott, & Berninger, 2007). If variability in effect size is due to definitional criteria used, then it is important to examine which definitional criteria is being employed by each study. This will give us a greater understanding of whether effect sizes found are a product of the definitional criteria used or the actual RD.

Word reading compared to comprehension

Research has identified children who have specific problems with reading comprehension despite adequate word reading skills (see Nation, 2007, for a review) and the pattern of RD seen in poor comprehenders is noted as being different from that seen in children with word reading problems (Hulme & Snowling, 2009). Executive functions, and in particular working memory, have been implicated in both disorders (Swanson, 2006; Vukovic & Siegel, 2006). A recent study by Sesma *et al.* (2009) found that executive functions contributed to reading comprehension ability, even after factors such as decoding, reading fluency, and vocabulary had been controlled for. Given that word RD and reading comprehension difficulties manifest as differing disorders, it is possible that there may be differences in executive function performance between children who have word reading difficulties (RD-WR) and those who have poor comprehension ability but good word reading skills (RD-RC). However, a study by Catts, Adolf, and Ellis Weismer (2006) compared children with RD-WR to those with RD-RC and control children. While the RD-RC were found to perform more poorly than controls on distance inference tasks, there was no significant difference between the RD-RC and RD-WR. The authors conceded that the distance inference task could be interpreted as evidence of working memory difficulties. This would imply that there may be a similar pattern of performance on tasks of working memory between these groups. It is therefore important to investigate whether the pattern of performance on tasks of executive function is the same for RD-WR and RD-RC or not and so this will be investigated in the present meta-analysis.

Age and gender

Previous research has shown that performance on tasks of executive function is influenced by both age (Davidson, Amso, Anderson, & Diamond, 2006) and gender (Lezak, 1995; cited in Reiter, Tucha, & Lange, 2005). In regard to the influence of age, Anderson, Anderson, Jacobs, and Smith (2008) reviewed the literature concerning the link between the development of executive functions and brain development throughout childhood. The review concluded that it would be expected that performance on tasks of executive function increase in line with brain development and thus be influenced by age. In addition, Giedd *et al.* (1996) highlighted that there are gender specific differences in brain development and it has been proposed that these differences may be related to hormone production (De Bellis *et al.*, 2001). It is therefore

plausible to suggest that there may be gender differences in performance on tasks of executive function. However, such an influencing role of gender and age on executive function task performance is contrary to some of the findings in the literature (e.g. Jerman & Swanson, 2005; O'Shaughnessy & Swanson, 1998) therefore the possible moderating influence of age and gender will be explored in the current meta-analysis.

Response modality

A review of the literature pertaining to immediate memory in children with learning disabilities in reading was carried out by O'Shaughnessy and Swanson (1998). A total of 41 studies were included and the overall standardized effect size was moderate (Cohen, 1988) in favour of children without reading problems (0.61, $SD = 0.87$), showing that children with RD show deficits in immediate memory. When moderator variables were explored the most prominent finding was that the RD group were most deficient on memory tasks which involved verbal material as opposed to visuospatial material.

In regard to working memory, which is thought to be an important aspect of executive function (Barkley, 1997; Denckla, 1996; Diamond *et al.*, 2007; Pennington *et al.*, 1996; Roberts & Pennington, 1996), Jerman and Swanson (2005) reviewed 28 studies of working memory in children with RD. These studies yielded an overall large mean effect size of 0.89 ($SE = .08$) thus indicating that children with RD are impaired on tasks assessing working memory. Age, IQ, reading level, and modality of the measures were not found to predict effect sizes which is in contrast to modality differences found for short-term memory tasks by O'Shaughnessy and Swanson (1998). This therefore suggests that the working memory deficit of RD is not restricted to verbal based tasks.

However, Nation, Adams, Bowyer-Crane, and Snowling (1999) investigated the working memory profile of a sample of children with deficits in reading and language comprehension and found that the performance of RD-RC was worse than TDC on tasks that involved verbal material but not on tasks that were non-verbal in nature. Given the discrepancies in the literature regarding working memory, it is possible that task modality may influence the magnitude of effect sizes found on tasks assessing all aspects of executive function for children with RD. Task modality will therefore be investigated as a possible moderator variable.

The present study: Research questions

As the literature reviewed above indicates, executive functions have been shown to be important in reading and to be impaired in children with RD. However, there is confusion in the literature arising from the range of measurement tasks used for assessment in the absence of a clear consensus regarding which aspects of executive function are measured by which task. It is therefore necessary to synthesize the literature to give an indication of which tasks of executive function consistently differentiate between RD and control groups and whether this pattern in this same for all measurement tasks.

To our knowledge, this is the first review of the performance of children with RD on tasks of executive function in general, rather than specific areas i.e. working memory. Based on the previous literature, the present meta-analysis addresses the following research questions:

(1) Which tasks of executive function discriminate best between children with RD and their typically developing peers? Are there differences in effect size depending on the task employed?
(2) Is the same pattern evident for children with IQ-discrepant RD and children with non-discrepant RD?
(3) Is the magnitude of the effects found for children with RD-WR different from the magnitude of effects for children who are RD-RC?
(4) What influence do age and gender of participants have on the magnitude of effect found?
(5) Does the response modality of tasks influence the results?

Method

Locating studies

A search was conducted of published studies examining executive functioning in school-age children with RD. Web-based search engines (Psych-info, Wilson Web, Web of Knowledge, and Pub-med) were used in order to locate papers published in peer-reviewed journals from 1974 until January 2008. Specific terms such as 'inhibition', 'inhibitory skills', 'executive function', and 'working memory' were entered in conjunction with terms such as 'reading', 'RD', 'reading disability', 'dyslexia', and 'children'. Appendix A shows the number of studies located in the varying search engines by keywords used. Reference citations from published studies were also consulted and authors currently active in the area were contacted.

Eligibility criteria

Eligibility criteria employed were (i) mean age of participants less than 16 years old; (ii) RD operationally defined as standard scores below 85 on a norm-referenced measure of reading ability or as a significant discrepancy between chronological age/ability and reading age; and (iii) descriptive or inferential statistics necessary to permit the calculation of effect sizes in regard to a comparison of executive function between children with RD who had no reported comorbidity and their typically developing peers. Appendix B provides details of studies which fulfilled these criteria and were therefore included in the meta-analysis and method of their retrieval.

Coding

Study coding

Sample demographics were retrieved from all studies which met the inclusion criteria. Information regarding the sample size, age, gender, and non-verbal IQ of participants was recorded from the Method section of studies which gave this information. Studies were also coded according to the criteria that had been used to determine whether participants had difficulties with reading, i.e. whether a discrepancy between IQ and reading attainment had been used or not.

The tasks used to measure executive functioning in each study were also recorded. Several studies which met the eligibility criteria had included tasks in their test battery used to assess abilities other than executive functioning. For example, some studies also

included tasks assessing short-term memory. As the present meta-analysis aimed only to investigate executive functions, data relating to these tasks were not included.

In addition, it was found that different names existed for the same task of executive function, for example, the commonly used backward digit span task had been labelled the numbers reversed in McGee *et al.* (2004). In these instances, the task was recorded as being the task under which it is most commonly known in order to allow ease of comparison, but only if the task followed the exact procedure. Where procedure or materials were different, the original name from the study was used.

Tasks were also coded based on whether they specifically required a verbal/language based or a non-verbal response. All coding was performed by the first author. Two independent coders were then trained in the coding procedure and, based on the task description given in the study from which the task was retrieved, coded 10% of the total number of tasks. The two independent coders had 100% agreement (Cohen's $\kappa = 1$).

Calculation of effect sizes

Where multiple outcomes were given for the same task, for example reaction time and number of errors made, the most common metric across studies was used to calculate the effect size. For example, if the majority of studies used number of errors as the outcome measure but one study reported both errors and reaction time, then number of errors was used when calculating the mean effect size. For tasks identified only once within the literature but with several outcome scores, the effect size was based upon the score which best discriminated the group of poor readers from the TDC.

Hedges' *g* standardized effect sizes with weight for sample size were calculated. Using this method, a positive effect size indicates better performance by the TDC.

Meta-analysis (procedure)

In total, 48 studies were located which fulfilled the eligibility criteria and from these 84 different tasks assessing executive function were identified which yielded 180 effect sizes (see Table 1 for study characteristics, tasks, and effect sizes). A meta-analysis of standardized effect sizes weighted for sample size was carried out using Comprehensive Meta Analysis (version 2) software (Borenstein, Hedges, Higgins, & Rothstein, 2005). One of the requirements for meta-analysis stated by Rosenthal (1994) is that it should not contain more than one effect size for each study included. However, the majority of studies identified that met with the criteria for the current meta-analysis contained several outcome measures, thus violating this assumption. While it is possible to perform a correlation of effect sizes and thus reduce the number of effect sizes to one per study, this would be counter to the objectives of the meta-analysis; that is, to investigate the array of tasks commonly employed. Therefore, to deal with this issue, meta-analysis was carried out on the lowest effect sizes from each study to provide the most conservative estimate, the lower bound analysis, and then rerun for the largest, the upper bound analysis (Law, Boyle, Harris, Harkness, & Nye, 1998). Details of tasks included in upper bound and lower bound analyses by modality and definitional criteria can be found in Appendix C. Following recommendations by the National Research Council (1992), instead of reporting the fixed-effects model, the more conservative random effects model is reported throughout.

Regression analyses were carried out to assess whether age, gender, and IQ of participants were significant predictors of effect sizes. Moderator variables included in

Table 1. Study characteristics: Sample size (number of males), age, non-verbal IQ, selection criteria, task- and bias-corrected effect size (Hedges' g), and standard error of effect size estimate (in parentheses)

Study	Sample characteristics						Task	Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria			
Altemeier et al. (2007)[a]	80 (80)	56 (56)	133	—	—	Dis	Inhibition/switching task	0.56 (.18)	3.11**
							Rapid automatic shifting	0.96 (.18)	5.33***
							Stroop	0.74 (.18)	4.11***
Bayliss, Jarrold, and Baddeley (2005)	50	50	125	—	—	Non-dis	Verbal working memory task	0.10 (.20)	0.50
							Visuospatial working memory task	0.23 (.20)	1.15
Authors (under revision)	29 (29)	24 (24)	126	—	—	Non-dis	Numerical Stroop (number detection)	0.52 (.28)	1.86
							Tower of London	0.42 (.28)	1.50
Brosnan et al. (2002)	30 (30)	30 (30)	167	—	—	Unsure	Group Embedded Figures Test	1.15 (.28)	4.11***
	16 (10)	16	119				Recognition	0.49 (.36)	1.36
	16 (10)	16	119				Temporal order	0.65 (.36)	1.81
	16 (10)	16	119				Verbal fluency	0.99 (.37)	2.68**
Cain and Oakhill (2006)	23 (8) +	23 (10)	91	9.85	11.15	Non-dis	Digit reading task	0.15 (.30)	0.50
							Listening span	0.71 (.30)	2.37*
Cain (2006)	13 (6) +	13 (4)	116	—	—	Non-dis	Word recall intrusion errors	1.32 (.43)	3.07**
Carretti et al. (2005)	109 (59) +	109 (59)	96–132	—	—	Non-dis	Updating task (delayed intrusions)	1.11 (.15)	7.40***
Censabella and Noel (2005)	15 (8)	15 (7)	136	—	—	Unsure	Backward digit span	1.06 (.39)	2.72**
							Flanker digits	1.18 (.40)	2.95**
							Flanker letters	1.57 (.42)	3.74***

Q4

Q3

Table 1. (*Continued*)

Study	Sample characteristics							Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria	Task		
Condor, Anderson, and Saling (1995)	18 (16)	18 (16)	107	112.50	108.70	Unsure	Numerical Stroop	1.24 (.40)	3.10**
							Stroop	0.63 (.37)	1.70
De Berni and Palladino (2000)	12 (7) +	12 (7)	101	8.75	9.33	Non-dis	Tower of Hanoi	0.72 (.34)	2.12*
							Backward digit span	0.22 (.41)	0.54
De Jong (1998)	18 (14)	18 (14)	123	–	–	Unsure	Listening span-intrusion	0.92 (.43)	2.14*
							Counting span	1.55 (.38)	4.08***
							Computation span	0.45 (.34)	1.32
							Reading span	1.62 (.38)	4.26***
							Star Counting Test	1.39 (.37)	3.76***
Everatt et al. (1997)	20 (20)	20 (20)	128	85.00	85.00	Unsure	Stroop	1.61 (.36)	4.43***
Everatt, Weeks, and Brooks (2008)	20 (14)	40 (25)	139	29.40	29.35	Unsure	Object interference (Stroop)	−0.09 (.27)	−0.33
							Spatial memory	−0.32 (.28)	−1.14
							Stroop	0.60 (.28)	2.14*
							Verbal span	0.71 (.28)	2.54*
Hall, Halperin, Schwartz, and Newcorn (1997)	17	28	116	94.90	96.50	Non-dis	Continuous performance task	0.54 (.31)	1.74
Helland and Asbjornsen (2000)	43b (36)	20 (16)	149	–	–	Unsure	Stroop	1.20 (.30)	4.00***
Howes, Bigler, Lawson, and Burlingame (1999)	42	20					Wisconsin Card Sorting	0.41 (.27)	1.52
	24 (16)	45 (30)	125	–	–	Unsure	Backward digit span	0.61 (.26)	2.35*
							Backward letter span	0.83 (.26)	3.19**
							Facial memory	0.24 (.25)	0.96
							Abstract visual memory	0.45 (.26)	1.73

Table 1. (*Continued*)

Study	Sample characteristics						Task	Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria			
Jeffries and Everatt (2004)	21 (18)	40 (35)	131	–	–	Non-dis	Backward digit span	1.48 (.30)	4.93***
							Listening span	0.43 (.27)	1.59
							Stroop	0.04 (.27)	0.15
							Verbal numerical Stroop	0.73 (.28)	2.61**
Kramer, Knee, and Delis (2000)	57 (44)	114 (88)	113	97.60	–	Dis	California Verbal Learning Test	0.18 (-.16)	1.13
Kupietz (1990)	11 (9)	11 (8)	116	–	–	Dis	Continuous performance task	1.05 (.45)	2.33*
Martinussen and Tannock (2006)	14 (9)	25 (17)	110	–	–	Non-dis	Backward digit span	1.44 (.37)	3.89***
	15	30					Reverse finger windows task	1.07 (.34)	3.15**
McGee et al. (2004)	28 (25)	42 (34)	106	–	–	Non-dis	Backward digit span	0.29 (.25)	1.16
McGee, Williams, Moffitt, and Anderson (1989)	13 (13)	62 (62)	156	100.10	107.90	Non-dis	Continuous performance	0.05 (.24)	0.21
							Delayed recall	0.03 (.31)	0.10
Miller-Shaul (2005)	25	25	~108	–	–	Non-dis	Verbal fluency	0.31 (.31)	1.00
							Trail making task	0.35 (.31)	1.13
							Wisconsin Card Sorting Task	0.06 (.31)	0.19
							Coding task	6.03 (.67)	9.00***
							Letter naming	1.52 (.32)	4.75***
							Number naming	1.30 (.29)	4.48***
							Object naming	0.68 (.29)	2.35*
							Verbal fluency	1.29 (.31)	4.16***
Närhi and Ahonen (1995)	21 (21)	10 (10)	128	92.30	86.50	Dis	Executive task (combined WCST errors and trail making task)	0.45 (.39)	1.15

Table 1. (*Continued*)

Study	Sample characteristics						Task	Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria			
Nation et al. (1999)	14 +	15	129	101.35	103.69	Non-dis	Naming	1.05 (.41)	2.56*
							Listening span	1.38 (.41)	3.37**
Nyden et al. (1999)	10 (10)	10 (10)	121	99.50	–	Dis	Spatial span	0.01 (.37)	0.03
							Arithmetic task	0.76 (.46)	1.65
							Coding task	0.81 (.47)	1.72
							Conflict task	0.87 (.47)	1.85
							Go/NoGo task	0.73 (.46)	1.59
							Symbol search	0.80 (.46)	1.74
							Wisconsin Card Sorting Task	0.17 (.45)	0.38
Pennington, Groisser, and Welsh (1993)	15 (15)	23 (23)	107	109.00	108.40	Dis	Continuous performance task	0.42 (.34)	1.24
Pickering and Gathercole (2004)	29 (21)	636	116.97m (pr)	–	–	Unsure	Matching Family Figures Test	–0.03 (.33)	–0.09
							Tower of Hanoi	0.08 (.33)	0.24
							Wisconsin Card Sorting Task	0.14 (.33)	0.42
							Backward digit span	0.32 (.19)	1.68
Protopapas, Archonti, and Skaloumbakas (2007)	16 (12)	72 (35)	150	33.60	40.00	Unsure	Counting span	0.14 (.19)	0.74
							Listening span	0.07 (.19)	0.37
							Stroop	0.86 (.30)	2.87**
Purvis and Tannock (2000)	17 (8)	17 (11)	113	12.00	12.30	Non-dis	Continuous performance task	–0.31 (.35)	–0.89
Reiter et al. (2005)	42 (26)	42 (26)	128	–	–	Dis	Stop-signal task	0.41 (.35)	1.17
							Animal test	0.43 (.23)	1.87
							Backward digit span	0.24 (.22)	1.09
							Five-point test	0.54 (.23)	2.35*

Table 1. (*Continued*)

Study	Sample characteristics						Task	Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria			
Roodenrys, Koloski, and Grainger (2001)	16	16	109	–	–	Unsure	Flexibility task	0.15 (.22)	0.68
							Go/NoGo task	0.16 (.22)	0.73
							Stroop	0.31 (.22)	1.41
							S-word test	0.24 (.22)	1.09
							Tower of London	−0.13 (.22)	−0.59
							Trail making task	0.25 (.22)	1.14
							Wisconsin Card Sorting Task	0.24 (.22)	1.09
							Auditory serial addition task	1.41 (.40)	3.53***
Savage and Frederickson (2006)	34 (22)	33 (24)	127	41.51	53.07	Non-dis	Memory updating (2 updates)	0.94 (.37)	2.54*
							Number generation (2 s/item)	0.43 (.36)	1.19
							Backward digit span	1.00 (.26)	3.85***
Stothard and Hulme (1992)	14 (7) +	14 (7)	95	–	–	Non-dis	Digit naming	0.80 (.25)	3.20**
							Object naming	0.19 (.24)	0.79
							Listening span	0.55 (.39)	1.41
Swanson (1993)	28 (20)	38 (30)	121	–	–	Non-dis	Auditory digit sequence	0.23 (.25)	0.92
							Mapping	0.38 (.25)	1.52
							Matrix	0.62 (.25)	2.48*
							Non-verbal sequencing	0.29 (.25)	1.16
							Phrase sequence	0.22 (.25)	0.88
							Picture sequence	0.27 (.25)	1.08
							Rhyming task	0.34 (.25)	1.36
							Semantic association	0.85 (.26)	3.27**
							Semantic categorization	0.48 (.25)	1.92

Table 1. (*Continued*)

Study	Sample characteristics RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria	Task	Hedges' g (SE)	Z value
Swanson (1999)	18 (17)	18 (9)	137	27.77	41.00	Unsure	Spatial organization	0.33 (.25)	1.32
							Story recall	0.67 (.26)	2.58*
							Counting span	1.20 (.36)	3.33**
							Matrix	1.27 (.37)	3.43**
							Sentence span	1.46 (.38)	3.84***
Swanson and Alexander (1997)	40 (32)	40 (24)	120	26.97	30.75	Unsure	Concurrent digit colour	0.72 (.23)	3.13**
							Concurrent digit semantic	0.72 (.23)	3.13**
							Concurrent digit shape	0.63 (.23)	2.74**
							Counting span	0.49 (.23)	2.13*
							Matrix	0.46 (.23)	2.00*
							Sentence span	0.68 (.23)	2.96**
Swanson and Ashbaker (2000)	30 (17)	30 (17)	179	–	–	Unsure	Auditory digit sequence	0.01 (.26)	0.04
							Mapping	1.25 (.28)	4.46***
							Matrix	0.88 (.27)	3.26**
							Sentence span	2.68 (.36)	7.44***
Swanson, Ashbaker, and Lee (1996)	20 (17)	34 (27)	150	–	–	Unsure	Auditory digit sequence	0.84 (.29)	2.90**
	20 (17)	34 (27)	150				Mapping	0.39 (.28)	1.39
	20 (17)	34 (27)	150				Matrix	–0.28 (.28)	–1.00
	30 (22)	30 (27)	174				Sentence span	1.48 (.29)	5.10***
	20 (17)	34 (27)	150				Story recall	0.85 (.29)	2.93**
Swanson and Berninger (1995)	22 (12)	33 (18)	160	–	–	Non-dis	Auditory digit sequence	0.76 (.28)	2.71**
							Mapping	0.53 (.28)	1.89
							Matrix	0.22 (.28)	0.79

Table 1. (*Continued*)

Study	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria	Task	Hedges' g (SE)	Z value
Swanson, Howard, and Saez (2006)	19	15	150	105.46	104.43	Unsure	Story recall	1.76 (.32)	5.50***
							Auditory digit sequence	0.93 (.36)	2.58*
							Backward digit span	0.83 (.36)	2.31*
							Coding task	0.35 (.35)	1.00
							Digit naming	1.45 (.39)	3.72***
							Letter generation	0.52 (.35)	1.49
							Letter naming	1.62 (.40)	4.05***
							Listening sentence span	0.16 (.35)	0.46
							Number generation	0.57 (.35)	1.63
							Object naming	1.39 (.38)	3.66***
							Semantic association task	0.81 (.36)	2.25*
							Updating task	0.41 (.35)	1.17
Swanson and Jerman (2007)	18	23	130	107.66	107.52	Unsure	Auditory digit sequence	0.75 (.33)	2.27*
							Backward digit span	1.19 (.34)	3.50***
							Rhyming task	1.36 (.35)	3.89***
							Updating task	0.64 (.32)	2.00*
Swanson et al. (2004)	31	61	75	105.55	103.45	Unsure	Backward digit span	0.17 (.22)	0.77
							Letter generation	−0.17 (.22)	−0.77
							Number generation	0.05 (.22)	0.23
							Rhyming task	0.20 (.22)	0.91
							Semantic association task	0.00 (.22)	0.00
Van der Schoot et al. (2004)	16 (9)	16 (6)	140	–	–	Unsure	Sentential priming task (% errors)	1.35 (.39)	3.46**
Van der Sluis et al. (2004)	21 (12)	19 (9)	129	–	–	Non-dis	Digit naming	0.34 (.32)	1.06
							Letter naming	0.79 (.33)	2.39*
							Object inhibition task	0.44 (.32)	1.38

Table 1. (*Continued*)

Study	Sample characteristics						Task	Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria			
							Object inhibition-shifting task	0.62 (.32)	1.94
							Object naming	0.52 (.32)	1.63
							Object shifting task	0.81 (.33)	2.46*
							Quantity inhibition task	0.15 (.32)	0.47
							Quantity naming task	0.17 (.32)	0.53
							Trail making task	0.08 (.32)	0.25
Van der Sluis, van der Leij, and de Jong (2005)	18 (11)	24 (11)	129	114.65	111.73	Unsure	Backward digit span	0.07 (.31)	0.23
Willcutt et al. (2001)	93	121	126	101.90	109.40	Dis	Listening span	0.19 (.31)	0.61
	75	102					Contingency Naming Test	0.28 (.14)	2.00*
	53	84					Continuous performance task	0.30 (.15)	2.00*
	53	84					Counting span	0.36 (.18)	2.00*
	75	102					Sentence span	0.36 (.18)	2.00*
	53	84					Stop-signal task	0.13 (.15)	0.87
	53	84					Stroop	0.31 (.14)	2.21*
	93	121					Trail making task	0.24 (.18)	1.33
							Wisconsin Card Sorting Task	0.28 (.14)	2.00*
Willcutt et al. (2005)	109 (56)	151 (65)	135	99.00	111.10	Non-dis	Arithmetic task	0.63 (.17)	3.71***
							Backward digit span	0.53 (.17)	3.12**
							Coding task	0.57 (.17)	3.35**
							Continuous performance task	0.32 (.13)	2.46*
							Counting span	0.46 (.18)	2.56*

Table 1. (*Continued*)

Study	Sample characteristics						Task	Hedges' g (SE)	Z value
	RD (n)	Control (n)	Age	RD NV-IQ	C NV-IQ	Criteria			
							Sentence span	0.47 (.13)	3.62***
							Spatial working memory task	0.29 (.13)	2.23*
							Stop-signal task	0.33 (.13)	2.54*
							Stroop	0.45 (.13)	3.46**
							Symbol search	0.67 (.13)	5.15***
							Trail making task	0.34 (.13)	2.62**
							Wisconsin Card Sorting test	0.22 (.13)	1.69

Note. RD, reading difficulties group; +, RD-RC; *p < .05; **p < .01; ***p < .001; Criteria: Dis, discrepancy criteria used to determine RD; Non-dis, Non-discrepancy criteria used to define RD; Unsure: no sufficient information provided to determine which criteria used.
[a] Dyslexic sample mean = 122 but subsample used. Control sample mean = 106 (56 male) in first year of cohort – this data taken from third year cohort in which 7 participants less, however no info given regarding gender of those who left. ES given based on original male sample, but also calculated with 7 less: RAS = 0.94 (.19), Stroop = 0.74 (.19), no change for inhibition/switching task.
[b] Data missing for 7 participants in RD group for Stroop task so n = 36.

the analysis were whether IQ/attainment discrepancy based criteria had been used to determine whether participants had RD and whether tasks required a verbal or non-verbal response.

Results

Description of studies included

The median number of participants per study was 52 (range 20–665). The age of participants ranged from 75 to 179 months with a median age of 126 months. The mean number of participants per study in the RD group was 29 ($SD = 24$) and in the control group was 49 ($SD = 92$). A total of 38 studies reported gender of participants and in the average study 74% of participants in the RD group were male and 70% of the control groups. A total of 23 studies provided specific information concerning the non-verbal ability of participants. However, only 15 studies provided information using a comparable metric, that is, figures for Full scale IQ were provided as opposed to raw scores for example, which could not be meaningfully compared. Additionally, this information was only provided for the control groups in 13 of the 15 studies. The mean non-verbal IQ of the RD group was 101.76 ($SD = 7.78$, range $= 85 – 114.65$, $n = 15$) and for the TDC was 103.41 ($SD = 8.79$, range 85–111.73, $n = 13$).

Publication bias

Effect sizes were plotted against standard errors to give a funnel plot as illustrated in Figure 1, indicating the presence of publication bias which was confirmed by significant results from Egger's test (Egger, Davey Smith, Schneider, & Minder, 1997) and Begg and Mazumdar's (1994) rank correlation (Kendall's tau-b). Rosenthal's (1979) Fail-safe N was 6,367, indicating that over 6,000 studies would be needed for the cumulative effect to be non-significant. Similarly, Orwin's (1983) Fail-safe N was 171, suggesting that more studies than included in the meta-analysis overall would need to be identified with a effect size of 0.00, before the cumulative effect would be 0.10, that is a small effect. Therefore, it can be concluded that the following meta-analysis provides a satisfactory representation of the relationship between executive functions and RD.
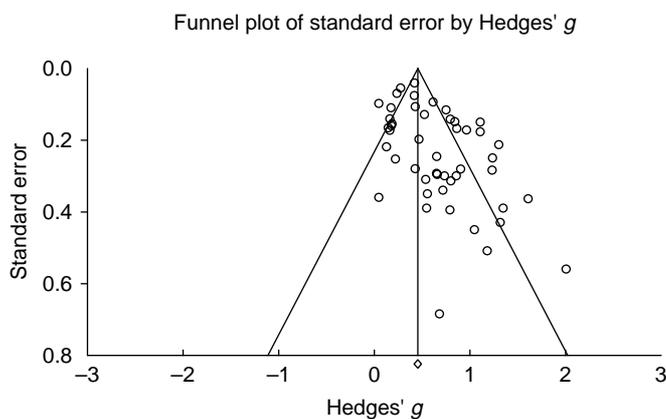
Figure 1. Funnel plot of effect size by standard error.

Overall analysis

Q4 *Z* scores were calculated as a test of the null-hypothesis and 91 of the 180 effect sizes calculated (50.56%) were significant. Effect sizes ranged from −0.32 (*SE* .28) to +1.83 (*SE* .84) with the overall mean weighted effect size being +0.57 (*SE* .03) in favour of the control groups. There was significant heterogeneity ($Q = 528.62$, $df = 179$, $p < .001$) and a moderate-to-large percentage of the variation was due to heterogeneity rather than chance ($I^2 = 66.14$), indicating that exploration of moderator variables was warranted. In order to ensure that the overall effect size was not a product of undue influence of one study, each effect size was removed in turn. With the removal of each study, effect sizes ranged from +0.55 to +0.57 and there continued to be significant heterogeneity. This sensitivity analysis therefore revealed that no single study had an undue influence on the overall effect size. Table 2 provides details of the different measurement tasks identified including effect sizes.

Inspection of Table 2 reveals that the coding subtest from the Wechsler Intelligence Scale for Children – Fourth Edition (WISC IV; Wechsler, 2004) was found to be the best at discriminating RD from their typically developing peers. This task was identified four times in the included papers and was associated with a mean effect size of +1.83 (*SE* .84; although see Discussion section for fuller interpretation). This task involves transcribing a digit-symbol code as quickly as possible and while it constitutes part of the processing speed factor derived in factor analytic studies of the WISC (see manual; Cockshott, Marsh, & Hine, 2006), it involves a variety of skills including attention and impulsivity (Flanagan & Kaufman, 2004; Nyden, Gillberg, Hjelmquist, & Heiman, 1999).

The least efficient task at discriminating RD from TDC was Nation *et al.*'s (1999) spatial span task which was only identified once in the review of literature, and was associated with an effect size of only 0.01 (*SE* .37). Descriptions of each task identified can be found in Appendix D.

Overall lower bound analysis

The overall mean weighted effect size for the lower bound analysis was +0.35 (*SE* .06). There was significant heterogeneity ($Q = 115.02$, $df = 47$, $p < .001$) and a moderate percentage of the variation in effect sizes was due to heterogeneity rather than chance ($I^2 = 59.14$). This suggested that the effects of moderator variables should be examined.

Overall upper bound analysis

The mean weighted effect size for the upper bound analysis was +0.97 (*SE* .09). Significant heterogeneity was found ($Q = 211.47$, $df = 47$, $p < .001$) and a moderate to large percentage of the variation of effect sizes was due to heterogeneity rather than chance ($I^2 = 77.77$). Thus, moderator variables were also examined for the upper bound analysis.

Moderator analysis

IQ/achievement discrepancy

Lower bound.   A smaller mean effect size was found when a discrepancy based criteria was employed (+0.24, *SE* .11) as opposed to a non-discrepancy criteria (+0.35, *SE* .10), however this difference failed to achieve statistical significance ($p > .05$).

Table 2. Details of different measurement tasks identified including mean bias corrected effect sizes (Hedges' g) and standard error of effect size estimate (in parentheses) (84 tasks)

Task (and number of exemplars)	Modality	Mean effect size (SE)	Z value
Abstract visual memory (1)	Non-verbal	0.45 (.26)	1.73
Animal test (1)	Verbal	0.43 (.23)	1.87
Arithmetic task (2)	Verbal	0.65 (.16)	4.05***
Auditory digit sequence (6)	Verbal	0.55 (.16)	3.46**
Auditory serial addition task (1)	Verbal	1.41 (.40)	3.53***
Backward digit span (14)	Verbal	0.63 (.12)	5.34***
Backward letter span (1)	Verbal	0.83 (.26)	3.19**
California Verbal Learning Test (1)	Verbal	0.18 (.16)	1.13
Coding task (4)	Non-verbal	1.83 (.84)	2.18*
Computation span (1)	Verbal	0.45 (.34)	1.32
Concurrent digit colour (1)	Non-verbal	0.72 (.23)	3.13**
Concurrent digit semantic (1)	Verbal	0.72 (.23)	3.13**
Concurrent digit shape (1)	Verbal	0.63 (.23)	2.74**
Conflict task (1)	Non-verbal	0.87 (.47)	1.85
Contingency Naming Test (1)	Verbal	0.28 (.14)	2.00*
Continuous performance task (7)	Non-verbal	0.29 (.10)	2.95**
Counting span (6)	Verbal	0.59 (.17)	3.52***
Delayed recall (1)	Verbal	0.03 (.31)	0.10
Digit naming (3)	Verbal	0.83 (.28)	2.91**
Digit reading task (1)	Verbal	0.15 (.30)	0.50
Executive task (1)	Non-verbal	0.45 (.39)	1.15
Facial memory (1)	Non-verbal	0.24 (.25)	0.96
Five-point test (1)	Non-verbal	0.54 (.23)	2.35*
Flanker digits (1)	Non-verbal	1.18 (.40)	2.95**
Flanker letters (1)	Non-verbal	1.57 (.42)	3.74***
Flexibility task (1)	Non-verbal	0.15 (.22)	0.68
Go/NoGo (2)	Non-verbal	0.30 (.25)	1.23
Group Embedded Figures Test (1)	Non-verbal	1.15 (.28)	4.11***
Inhibition/switching test (1)	Verbal	0.56 (.18)	3.11**
Letter generation (2)	Verbal	0.12 (.34)	0.36
Letter naming (3)	Verbal	1.29 (.27)	4.86***
Listening span – intrusions (1)	Verbal	0.92 (.43)	2.14*
Listening span (7)	Verbal	0.43 (.15)	2.81**
Mapping (4)	Non-verbal	0.63 (.20)	3.09**
Matching Family Figures Test (1)	Non-verbal	−0.03 (.33)	−0.09
Matrix (6)	Non-verbal	0.51 (.20)	2.57*
Memory updating (2 updates) (1)	Verbal	0.94 (.37)	2.54*
Naming (1)	Verbal	1.05 (.41)	2.56*
Non-verbal sequencing (1)	Non-verbal	0.29 (.25)	1.16
Number generation (3)	Verbal	0.25 (.17)	1.49
Number naming (1)	Verbal	1.30 (.29)	4.48***
Numerical Stroop (1)	Non-verbal	0.52 (.28)	1.86
Numerical Stroop (1)	Verbal	1.24 (.4)	3.10**
Object-inhibition-shifting task (1)	Verbal	0.62 (.32)	1.94
Object interference (Stroop) (1)	Verbal	−0.09 (.27)	−0.33
Object naming (4)	Verbal	0.64 (.24)	2.71**
Object shifting (1)	Verbal	0.81 (.33)	2.46*
Object-inhibition task (1)	Verbal	0.44 (.32)	1.38

Table 2. (*Continued*)

Task (and number of exemplars)	Modality	Mean effect size (*SE*)	Z value
Phrase sequence (1)	Verbal	0.22 (.25)	0.88
Picture sequence (1)	Non-verbal	0.27 (.25)	1.08
Quantity inhibition task (1)	Verbal	0.15 (.32)	0.47
Quantity naming task (1)	Verbal	0.17 (.32)	0.53
Rapid automatic shifting (1)	Verbal	0.96 (.18)	5.33***
Reading span (1)	Verbal	1.62 (.38)	4.26***
Recognition task (1)	Non-verbal	0.49 (.36)	1.36
Reverse finger windows task (1)	Non-verbal	1.07 (.34)	3.15**
Rhyming task (3)	Verbal	0.58 (.31)	1.85
Semantic association task (3)	Verbal	0.52 (.31)	1.70
Semantic categorization (1)	Verbal	0.48 (.25)	1.92
Sentence span (6)	Verbal	1.13 (.29)	3.86***
Sentential priming task (1)	Verbal	1.35 (.39)	3.46**
Spatial memory (1)	Non-verbal	−0.32 (.28)	−1.14
Spatial organization (1)	Non-verbal	0.33 (.25)	1.32
Spatial span (1)	Non-verbal	0.01 (.37)	0.03
Spatial working memory task (1)	Non-verbal	0.29 (.13)	2.23*
Star Counting Test (1)	Verbal	1.39 (.37)	3.76***
Stop signal (3)	Non-verbal	0.26 (.10)	2.71**
Story recall (3)	Verbal	1.07 (.32)	3.32**
Stroop (10)	Verbal	0.61 (.12)	5.19***
S-word test (1)	Verbal	0.24 (.22)	1.09
Symbol search (2)	Non-verbal	0.68 (.13)	5.43***
Temporal order (1)	Non-verbal	0.65 (.36)	1.81
Tower of Hanoi (2)	Non-verbal	0.40 (.32)	1.23
Tower of London (2)	Non-verbal	0.12 (.27)	0.43
Trail making task (5)	Non-verbal	0.28 (.09)	3.24**
Updating task (2)	Verbal	0.54 (.24)	2.27*
Updating task (delayed intrusion) (1)	Verbal	1.11 (.15)	7.40***
Verbal fluency (3)	Verbal	0.86 (.31)	2.80**
Verbal numerical Stroop (1)	Verbal	0.73 (.28)	2.61**
Verbal span (1)	Verbal	0.71 (.28)	2.54*
Verbal working memory task (1)	Verbal	0.10 (.20)	0.50
Visuospatial working memory task (1)	Non-verbal	0.23 (.20)	1.15
Wisconsin Card Sorting Task (7)	Non-verbal	0.24 (.08)	3.13**
Word recall intrusion errors (1)	Verbal	1.32 (.43)	3.07**

*p < .05; **p < .01; ***p < .001.

Upper bound.   When participants were selected based on a discrepancy between IQ and achievement a smaller mean effect size was found (+0.60, *SE* .13) than when a non-discrepancy criteria (+1.00, *SE* .16) was utilized. This difference was found to be statistically significant ($Q = 3.90$, $df = 1$, $p < .05$, $I^2 = 79.76$).

Is the magnitude of the effects found for children with RD-WR the same as for children who are poor comprehenders (RD-RC)?   Participants in six of the studies included were defined as being RD-RC namely studies by Cain (2006), Cain and Oakhill (2006),

Table 3. Summary of upper and lower bound results

	Overall	IQ discrepancy		Modality	
		Discrepant	Non-discrepant	Verbal	Non-verbal
Lower bound	+0.35 (0.06)	+0.24 (0.11)	+0.35 (0.10)	+0.45 (0.09)	+0.22 (0.07)*
Upper bound	+0.97 (0.09)	+0.60 (0.13)	+1.00 (0.16)*	+1.02 (0.10)	+0.89 (0.18)
Overall	+0.57 (0.03)				

*$p < .05$.

Carretti, Cornoldi, De Beni, and Romano (2005), De Berni and Palladino (2000), Nation *et al.* (1999), and Stothard and Hulme (1992). That is, participants had average word reading skills but their reading comprehension was significantly poorer. The remaining 42 studies employed participants with RD-WR whose main operational definition centred on word reading. Using a random effects analysis, the mean effect size for the RD-RC was +0.82 (*SE* .16) and for the participants with RD-WR was +0.59 (*SE* .05). This difference was not found to be statistically significant ($Q = 1.95$, $df = 1$, $p > .05$), which justifies treating the RD as one group for the purpose of this analysis.

Meta-regression analyses
Meta-regression analyses were carried out to assess whether the magnitude of effect found varied as a function of the sample characteristics. The mean effect size from each study was used as the criterion variable and the age of participants, percentage of males in the RD group, and non-verbal IQ of the RD group and of the control group were all employed as predictor variables. As not all studies reported the necessary information, regression analyses were performed on a subsample of studies which did provide the relevant information. None of these variables were found to significantly predict effect size (all p values > 0.05).

Response modality

Lower bound.   For the lower bound analysis, tasks which required a verbal response had a higher effect size (+0.45, *SE* .09) than tasks which required a non-verbal response (+0.22, *SE* .07). This difference was found to be statistically significant ($Q = 3.97$, $df = 1$, $p < .05$) and a moderate percentage of the variation of effect sizes was due to heterogeneity rather than chance ($I^2 = 59.14$).

Upper bound analysis.   The mean effect size for the tasks which required a verbal response in the upper bound analysis was +1.02 (*SE* .10) and for tasks which a required a non-verbal response was +0.89 (*SE* .18). This difference was not found to be statistically significant, however ($p > .05$).
The findings of the upper and lower bound analyses are summarized in Table 3.

Discussion
The results of the present meta-analysis found that children with RD are generally impaired in executive functioning when compared with TDC, with a medium-sized

effect (Cohen, 1988) being found overall in favour of TDC. However, the findings highlight that there is wide variation in the magnitude of effects found which relate to the assessment tasks utilized. Moderators such as task modality and the criteria used to define RD were found to moderate the magnitude of these effects. The findings therefore suggest that children with RD may have more pronounced difficulties in some areas of executive function compared to others.

Task discrimination

The task found to discriminate best between RD and TDC was the coding subtest from the WISC IV (Wechsler, 2004). This task is seen to involve a variety of skills including attention and impulsivity (Flanagan & Kaufman, 2004; Nyden *et al.*, 1999) however it also constitutes part of the processing speed factor derived in factor analytic studies of the WISC (see manual; Cockshott *et al.*, 2006). The different interpretations about what underlying constructs are measured by this task is a problem which is inherent to research on executive function. Poor performance on this task could be taken as evidence that children with RD have attentional difficulties but there may be other interpretations.

While this task was found to have the largest effect size, caution must be taken not to over-interpret these results. Four different effect sizes were generated by this task, however one study in particular (Miller-Shaul, 2005) contributed an extremely large effect size (+6.03) thus increasing the overall effect size somewhat. The participants in the Miller-Shaul study seem to be particularly impaired on the processes that the coding subtest is assessing, however, as this task was not a language based task the results cannot be attributed to any language-based differences between the participants in this sample and those in the other samples which also used this task. It must be noted however, that the participants in the Miller-Shaul study were Hebrew speakers (although see the limitations section for a fuller discussion).

When the effect size from the Miller-Shaul study was removed, the reading span task employed by De Jong (1998) became the task with the greatest effect size, however, this task was only identified once within the included studies, indicating that further research using this task is required before it can be concluded that it discriminates well between RD and controls.

Moderator analysis

Does IQ discrepancy have any utility?

The criteria used to select RD was evaluated as a moderator variable. In some studies, a discrepancy between IQ and attainment was used to define RD, while in others, no such discrepancy was utilized and RD were defined on low reading level alone. Differences in the magnitude of effect were identified but only in the upper bound analysis, with non-discrepancy criteria being associated with a higher mean effect size. As no significant differences were found in the lower bound analysis it could be that there are no fundamental differences in executive function profile between discrepant RD and non-discrepant RD and that actually the observed differences are a product of the assessment task utilized, as different tasks were identified in the upper and lower bound analyses with only a third of the tasks in the upper bound analysis also found in the lower bound.

This finding is supportive of the meta-analysis carried out by Stuebing *et al.* (2002), and is consistent with the position of Stanovich (2005) who continues to argue against

the use of IQ-discrepancy definitions of RD and suggests that IQ-discrepant readers do not differ from IQ-consistent readers on any of the propositions which would support such a distinction. The findings from the present meta-analysis indicate that statistically significant differences can be found depending on which definition is used, but only as a function of assessment task. This could be taken as evidence that tasks of executive function vary in how much they implicate fluid intelligence, and in fact that different aspects of executive functions vary in how strongly they are related to intelligence. However, Swanson (2006) discusses the literature which assesses the relationship between working memory and intelligence in children with RD. He concluded that children with RD are impaired on tasks of working memory even when intelligence is taken into account. This supports the premise that executive function impairments seen in children with RD are not simply a consequence of whether they have IQ discrepant RD or not.

However, as differences were only found in the upper bound analysis in the present study with possible implications for the utility of non-verbal IQ in defining RD – this conclusion cannot be unequivocally accepted. As the findings in the present meta-analysis were not consistent and instead varied as a function of task clear conclusions cannot yet be drawn and further research is needed to investigate whether the findings of Swanson (2006) are the case for all aspects of executive function or restricted to working memory. Until this complex underlying relationship is more fully understood however, care is required when selecting tasks of executive function and indeed, measures of more general cognitive functioning, which may also have implications for definitional criteria.

Word reading and reading comprehension
Six of the studies included in the present meta-analysis involved children whose RD were based on comprehension difficulties rather than word RD. Analyses found that the magnitude of effect found did not differ as a function of these group differences. This suggests that the executive function profile of children with RD-WR is no different from those with RD-RC. This is contrary to research by Cutting, Materek, Cole, Levine, and Mahone (2009) who found prominent difficulties on tasks of executive function for those with comprehension difficulties but not those with word RD. However, it must be acknowledged that the results of the present meta-analysis could be related to the unequal sample sizes, and so further research is needed in order to discern whether these groups truly do have the same pattern of results in terms of executive function task performance.

Age and gender
Regression analyses found that neither age nor gender of participants influenced the magnitude of effect sizes found. This finding is consistent with the findings of previous meta-analyses which have looked at working memory and RD (Jerman & Swanson, 2005; O'Shaughnessy & Swanson, 1998), however, the suggestion that age has no bearing on executive function performance does not sit comfortably within the developmental literature. One possible explanation could be that as the majority of studies in this meta-analysis involved participants aged between 114 months and 138 months, this age range is not wide enough to be sensitive to any age-related differences in performance. A further possible explanation is that different executive functions have

different developmental trajectories (Davidson *et al.*, 2006) and thus the array of executive functions involved in this meta-analysis concealed any developmental differences. Longitudinal research is thus required to fully understand the pattern of developmental changes.

Modality

Modality differences were found in general in the more conservative lower bound analysis, with verbal tasks producing a significantly larger effect than non-verbal tasks. While the extent of the difference in effect sizes between verbal and non-verbal tasks for the upper bound analysis was not found to be statistically significant, a large effect size was found for the tasks which required a non-verbal response as well as tasks which required a verbal response.

Mixed findings regarding the role of task modality have been reported in previous studies. O'Shaughnessy and Swanson (1998) found a domain general deficit on working memory tasks for children with RD-WR whereas Nation *et al.* (1999) reported a domain-specific deficit for RD-RC. Our findings suggest that children with RD have a depressed performance in general on tasks of executive function but that this depression increases as a function of increased language demands of the tasks. Given that there is approximately a 50% overlap between RD and specific language impairment (SLI; McArthur, Hogben, Edwards, Heath, & Mengler, 2000) it is of no surprise that the language demands of the tasks influence the performance of children with RD. However, research by Archibald and Gathercole (2006) found that children with SLI had impairments with tasks of working memory even when language age was used to calculate standard scores. Thus their deficient performance on tasks of working memory was above and beyond their difficulties with language. It therefore seems unlikely that the performance of children with RD is completely a function of any difficulties with language they may have. Furthermore, as large effect sizes were found for non-verbal tasks, it can be implied that RD have general impairments with executive function tasks, however these impairments will be more pronounced when the language demands of the tasks are increased.

This has many implications for the assessment of executive function in children with RD as it highlights that in order to gain a full understanding of the nature of the executive function impairment, non-verbal tasks should be emphasized (Booth *et al.*, in press). The contradictory results found for children with comprehension difficulties could reflect underlying differences between the causes of these impairments; however, as suggested previously, further research is needed to understand these potential group differences.

The relation to theoretical accounts

Theoretical accounts of reading imply the involvement of higher order cognitive processes such as activation and inhibition (Lupker, 2007) and research into RD suggests that executive function impairments may be integral (Swanson, 2006). While there continues to be discussion in the executive function literature regarding which executive function may be more dominant in general, several theories highlight the importance of both inhibition and working memory (Barkley, 1997; Denckla, 1996; Diamond *et al.*, 2007; Pennington *et al.*, 1996; Roberts & Pennington, 1996). The results of the present meta-analysis indicate that the executive function impairment of children

with RD is not uniform across all tasks, thus suggesting that some areas of executive function may be more impaired than others. However, until further research allows the exact nature of the executive function impairment to be isolated, it is not possible to say whether the impairment is specific to areas such as inhibition and working memory, or involves more areas, but perhaps just to differing degrees.

Limitations of the present study

One limitation of the present study concerns the lack of unpublished studies. While authors currently active in the area were contacted, only one unpublished study was identified. However, the analysis concerning publication bias revealed that the present study is an adequate representation of this area. A further limiting factor is that not all studies gave information concerning the non-verbal IQ of participants. In order to investigate the impact of non-verbal IQ, regression analysis were therefore carried out on the subsample of studies which did provide this information.

Furthermore, while studies included in the meta-analysis were confined to those whose participants had RD with no reported comorbidity, it is important to acknowledge that some of the participants in these studies may have had undiagnosed difficulties. These difficulties could influence their performance on tasks of executive function which would therefore be reflected in the results of this review.

In addition, it must be noted that some studies included in the present meta-analysis included participants who were not English speakers (i.e. the study by Miller-Shaul (2005) included Hebrew speakers). While a review of the nature of RD in different languages is beyond the scope of this paper, the considerable disagreement in the literature regarding the underlying causes of RD in different languages must be acknowledged. In a review of the literature, Caravolas (2007) states that findings are mixed; some researchers posit that differing writing systems influence the nature of the deficits seen in RD, whereas others argue that phonological skills play a role in RD despite differences in orthographic transparency. There seems to be no consensus at the present time, however it is also perhaps worth mentioning that both English and Hebrew can be considered deep orthographies (Frost, 2007). Nevertheless, the differing languages of participants could be considered a limitation of the present study.

One further limitation is that there are several outcome measures for studies in many cases. Combined with the issues of task impurity and in the absence of complete data sets, this resulted in the use of the upper and lower bound analyses. While the use of separate upper and lower bound analyses allowed for comparison of several effect sizes per study, it has the limitation of not allowing a mapping between tasks and the putative executive function which they measure, something which is further complicated by task impurity.

The fact that many executive tasks implicate several areas of functioning limits the conclusions which can be drawn from the present meta-analysis. This issue of task impurity means that it is not possible to argue that children with RD have impairments with some aspects of executive function but not with others as we can never be entirely sure that we are not assessing several aspects of executive function, but just to differing degrees. This is coupled with the fact that there are differing opinions about the underlying constructs which are measured by each task. The results of the present meta-analysis do highlight that RD do not have a uniform impairment with executive functioning though, although it is not possible to say exactly which areas are more impaired than others.

Research using several tasks of executive function and analysis using a latent variable approach will allow clearer conclusions to be drawn. Latent variable analysis determines the extent to which tasks implicate common constructs and also the degree of overlap between these constructs. Further studies assessing latent constructs and the extent that tasks load on to underlying theoretical constructs of executive function should therefore go some way to address these issues.

Implications

Burgess, Alderman, Evans, Emslie, and Wilson (1998, p. 556) argued that 'If different executive tasks measure different aspects of the dysexecutive syndrome, it makes sense to administer, standardly, a variety of tests rather than relying on just one or two'. The results of the present meta-analysis support this conclusion. While the mean effect size found for performance on tasks of executive function of RD compared to TDC was medium in effect, it ranged considerably, depending on the task and the underlying demands. This highlights the fact that results found can vary considerably depending on the assessment task utilized.

To take a practical example in regards to working memory, the Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001) is a widely used test battery within the UK which uses tasks which were identified in the present meta-analysis. The three tasks used to assess central executive function within this battery are the backward digit recall task, the counting recall task, and the listening recall task; all of which are included in the present meta-analysis. Mean effect sizes found for these individual tasks ranged from $+0.43$ to $+0.63$, and the mean effect size for this battery was found to be $+0.57$ (*SE* .08); that is, a medium effect. Furthermore, the mean effect size identified for participants who had been defined as RD using a non-discrepant definition was higher ($+0.72$, *SE* .13) than the mean effect size found using discrepancy criteria ($+0.31$, *SE* .14), a difference which was found to be statistically significant ($Q = 4.97$, $df = 1$, $p < .05$). Therefore, the WMTB-C can be seen to be sensitive to differences between RD and TDC, especially if non-discrepancy criteria are employed. While other working memory batteries may be equally sensitive to between-group differences, without being able to locate the tasks they use within those identified in this review, conclusions about their sensitivity cannot be drawn. This example serves to highlight the practical implications of assessing working memory in children with RD and the significance of using discrepancy criteria for the assessment of RD.

Ultimately, the tasks which are employed will depend upon the researcher's theoretical orientation regarding the underlying constructs that each task measures and also the research questions being addressed. However, the findings of the present meta-analysis may assist researchers in identifying appropriate tasks of executive function to maximize sensitivity of between-group comparisons, for example, between discrepant and non-discrepant RD.

Conclusions

Discrepancies have been identified in the literature regarding whether children with RD show impairments on tasks of executive functioning, even when the same test is used, for example, the Stroop test (e.g. Helland & Asbjornsen, 2000). The findings from the present meta-analysis indicate that children with RD do have impairments with

executive function and that the discrepancies found in the literature could be a product of the wide variety of assessment tasks being used and also that differences between the criteria used to select RD may account for further variation. Thus, it seems that both researchers and practitioners alike need to give their task selection considerable thought, not only towards which tasks help answer the research hypotheses but also including consideration of the underlying task demands and participant characteristics.

Acknowledgements

References

Asterisk indicates study included in meta-analysis.

*Altemeier, L. E., Abbott, R. D., & Berninger, V. W. (2007). Executive functions for reading and writing in typical literacy development and dyslexia. *Journal of Clinical and Experimental Neuropsychology*, *30*, 1-19.

Anderson, V., Anderson, P. J., Jacobs, R., & Smith, M. S. (2008). Development and assessment of executive function: From preschool to adolescence. In V. Anderson, R. Jacobs, & P. J. Anderson (Eds.), *Executive functions and the frontal lobes* (pp. 123-154). Hove: Psychology Press.

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders*, *41*(6), 673-693.

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

Baron, I. S. (2004). *Neuropsychological evaluation of the child*. Oxford: Oxford University Press.

*Bayliss, D. M., Jarrold, C., Baddeley, A. D., & Leigh, E. (2005). Differential constraints on the working memory and reading abilities of individuals with learning difficulties and typically developing children. *Journal of Experimental Child Psychology*, *92*, 76-99.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*, 1088-1101.

*Booth, J. N., Boyle, J. M. E., & Kelly, S. W. (in press). *The role of inhibitory functioning in children's reading skills*. Educational Psychology in Practice.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version2*. Englewood, NJ: Biostat.

*Brosnan, M., Demetre, J., Hamill, S., Robson, K., Shepherd, H., & Cody, G. (2002). Executive functioning in adults and children with developmental dyslexia. *Neuropsychologia*, *40*, 2144-2155.

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

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M.-A., Dawson, D. R., … Channon, S. (2006). The case for the development and use of 'ecologically valid' measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, *12*, 194-209.

*Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory*, *14*(5), 553-569.

*Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, *76*, 683-696.

Caravolas, M. (2007). The nature and causes of dyslexia in different languages. In M. J. Snowling & C. Hulme (Eds.), *The science of reading; a handbook* (pp. 336-355). Oxford: Blackwell Publishing.

Q6

Q7
Q8

Q9

*Carretti, B., Cornoldi, C., De Beni, R., & Romano, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology*, *91*, 45-66.

Catts, H. W., Adolf, S. M., & Ellis Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech Language and Hearing Research*, *49*, 278-293.

*Censabella, S., & Noel, M. P. (2005). The inhibition of exogenous distracting information in children with learning disabilities. *Journal of Learning Disability*, *38*(5), 400-410.

Cockshott, F. C., Marsh, N. V., & Hine, D. W. (2006). Confirmatory factor analysis of the Wechsler Intelligence Scale for Children, Third Edition in an Australian clinical sample. *Psychological Assessment*, *18*(3), 353-357.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Condor, A., Anderson, V., & Saling, M. (1995). Do reading disabled children have planning problems? *Developmental Neuropsychology*, *11*(4), 485-502.

Conners, C. K. (1995). *Conners' Continuous Performance Test users' manual*. Toronto: Multi-Health Systems.

Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia*, *59*, 34-54.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*(4), 422-433.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037-2078.

De Bellis, M. D., Keshavan, M. S., Beers, S. R., Hall, J., Frustaci, K., Masalehdan, A., … Boring, A. M. (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex*, *11*, 552-557.

*De Berni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks. Are they related to reading comprehension ability. *Learning and Individual Differences*, *12*, 131-143.

*De Jong, P. F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology*, *70*, 75-96.

Dempster, F. N., & Cooney, J. B. (1982). Individual differences in digit span, susceptibility to proactive interference, and aptitude/achievement test scores. *Intelligence*, *6*, 399-416.

Dempster, F. N., & Corkill, A. J. (1999). Individual differences in susceptibility to interference and general cognitive ability. *Acta Psychologica*, *101*, 395-416.

Denckla, M. B. (1996). A theory and model of executive function; A neuropsychological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 263-278). Baltimore, MD: Brookes.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*, 1387-1388.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629-634.

Ellis Weismer, S., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1249-1260.

*Everatt, J., Warner, J., Miles, T. R., & Thomson, M. E. (1997). The incidence of Stroop interference in dyslexia. *Dyslexia*, *3*, 222-228.

*Everatt, J., Weeks, S., & Brooks, P. (2008). Profiles of strengths and weaknesses in dyslexia and other learning difficulties. *Dyslexia*, *14*, 16-41.

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Wiley: Hoboken, NJ.

Frost, R. (2007). Orthographic systems and skilled word recognition processes in reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading; a handbook* (pp. 272–295). Oxford: Blackwell Publishing.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology, 70*, 177–194.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science, 4*, 294–298.

Giedd, J. N., Snell, J. W., Lange, N., Rajapakse, J. C., Casey, B. J., Kozuch, P. L., … Rapoport, J. L. (1996). Quantitative magnetic resonance imaging of human brain development: Ages 4–18. *Cerebral Cortex, 6*, 551–560.

*Hall, S. J., Halperin, J. M., Schwartz, S. T., & Newcorn, J. H. (1997). Behavioural and executive functions in children with attention-deficit hyperactivity disorder and reading disability. *Journal of Attention Disorders, 1*(4), 235–247.

*Helland, T., & Asbjornsen, A. (2000). Executive functions in dyslexia. *Child Neuropsychology, 6*(1), 37–48.

*Howes, N. L., Bigler, E. D., Lawson, J. S., & Burlingame, G. M. (1999). Reading disability subtypes and the Test of Memory and Learning. *Archives of Clinical Neuropsychology, 14*(3), 317–339.

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language, learning and cognition*. Oxford: Wiley-Blackwell.

*Jeffries, S., & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia, 10*, 196–214.

Jerman, O., & Swanson, H. L. (2005). Working memory and reading disabilities: A selective meta-analysis of the literature. *Advances in Learning and Behavioural Disabilities; Cognition and Learning in Diverse Settings, 18*, 1–31.

*Kramer, J. H., Knee, K., & Delis, D. C. (2000). Verbal memory impairments in dyslexia. *Archives of Clinical Neuropsychology, 15*(1), 83–93.

*Kupietz, S. S. (1990). Sustained attention in normal and in reading-disabled youngsters with and without ADDH. *Journal of Abnormal Child Psychology, 18*(4), 357–372.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (1998). Screening for speech and language delay: A systematic review of the literature. *Health Technology Assessment, 2*(9), 1–184.

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

Lupker, S. J. (2007). Visual word recognition: Theories and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading; a handbook* (pp. 39–60). Oxford: Blackwell Publishing.

*Martinussen, R., & Tannock, R. (2006). Working memory impairments in children with attention-deficit hyperactivity disorder with and without comorbid language learning disorders. *Journal of Clinical and Experimental Neuropsychology, 28*, 1073–1094.

McArthur, G. M., Hogben, J. H., Edwards, S. M., Heath, S. M., & Mengler, E. D. (2000). On the 'specifics' of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry, 41*(7), 869–874.

*McGee, R., Brodeur, D., Symons, D., Andrade, B., & Fahie, C. (2004). Time perception: Does it distinguish ADHD and RD children in a clinical sample? *Journal of Abnormal Child Psychology, 32*(5), 481–490.

*McGee, R., Williams, S., Moffitt, T., & Anderson, J. (1989). A comparison of 13-year-old boys with attention deficit and/or reading disorder on neuropsychological measures. *Journal of Abnormal Child Psychology, 17*(1), 37–53.

*Miller-Shaul, S. (2005). The characteristics of young and adult dyslexic readers on reading and reading related cognitive tasks as compared to normal readers. *Dyslexia, 11*, 132–151.

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

Morris, R. D. (1996). Relationships and distinctions among the concepts of attention, memory, and executive function: A developmental perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 11–16). Baltimore, MD: Brookes.

*Närhi, V., & Ahonen, T. (1995). Reading disability with or without attention deficit hyperactivity disorder: Do attentional problems make a difference? *Developmental Neuropsychology*, *11*(3), 337–349.

Nation, K. (2007). Children's reading comprehension difficulties. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248–266). Oxford: Blackwell Publishing.

*Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, *73*, 139–158.

National Research Council (1992). *Combining information: Statistical issues and opportunities for research*. Washington, DC: National Academy Press.

*Nyden, A., Gillberg, C., Hjelmquist, E., & Heiman, M. (1999). Executive function/attention deficits in boys with Asperger syndrome, attention disorder and reading/writing disorder. *Autism*, *3*(3), 213–228.

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational and Behavioural Statistics*, *8*, 157–159.

O'Shaughnessy, T. E., & Swanson, H. L. (1998). Do immediate memory deficits in students with learning disabilities in reading reflect a developmental lag or deficit?: A selective meta-analysis of the literature. *Learning Disabilities Quarterly*, *21*(2), 123–148.

Ozonoff, S. (1997). Components of executive function in autism and other disorders. In J. Russell (Ed.), *Autism as an executive disorder* (pp. 179–211). Oxford: Oxford University Press.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory and Cognition*, *29*(2), 344–354.

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

*Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity versus reading disability. *Developmental Psychology*, *29*(3), 511–523.

Pickering, S. J., & Gathercole, S. E. (2001). *Working Memory Test Battery for Children (WMTB-C)*. London: Psychological Corporation.

*Pickering, S. J., & Gathercole, S. E. (2004). Distinctive working memory profiles in children with special educational needs. *Educational Psychology*, *24*(3), 393–408.

*Protopapas, A., Archonti, A., & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology*, *54*, 251–282.

*Purvis, K. L., & Tannock, R. (2000). Phonological processing, not inhibitory control, differentiates ADHD and reading disability. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(4), 485–494.

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–38). Hove: Psychology Press.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hove: Erlbaum.

*Reiter, A., Tucha, O., & Lange, K. W. (2005). Executive functions in children with dyslexia. *Dyslexia*, *11*, 116–131.

*Roodenrys, S., Koloski, N., & Grainger, J. (2001). Working memory function in attention deficit hyperactivity disordered and reading disabled children. *British Journal of Developmental Psychology*, *19*, 325–337.

Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–281). New York: Russell Sage Foundation.

*Savage, R. S., & Frederickson, N. (2006). Beyond phonology: What else is needed to describe the problems of below-average readers and spellers? *Journal of Learning Disabilities*, *39*(5), 399–413.

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*(3), 232–246.

Stanovich, K. E. (2005). The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience? *Learning Disability Quarterly*, *28*(2), 103–106.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition and working memory. *Quarterly Journal of Experimental Child Psychology*, *59*(4), 745–759.

*Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing: An Interdisciplinary Journal*, *4*, 245–256.

Stuebing, K. L., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal*, *39*(2), 469–518.

*Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology*, *56*, 87–114.

*Swanson, H. L. (1999). Reading comprehension and working memory in learning disabled readers: Is the phonological loop more important than the executive system. *Journal of Experimental Child Psychology*, *72*, 1–31.

Swanson, H. L. (2006). Working memory and reading disabilities: Both phonological and executive processing deficits are important. In T. P. Alloway & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 59–83). Hove: Psychology Press.

*Swanson, H. L., & Alexander, J. E. (1997). Cognitive processes as predictors of word recognition and reading comprehension in learning-disabled and skilled readers: Revisiting the specificity hypothesis. *Journal of Educational Psychology*, *89*(1), 128–158.

*Swanson, H. L., & Ashbaker, M. H. (2000). Working memory, short-term memory, speech rate, word recognition and reading comprehension in learning disabled readers: Does the executive system have a role? *Intelligence*, *28*(1), 1–30.

*Swanson, H. L., Ashbaker, M. H., & Lee, C. (1996). Learning-disabled readers' working memory as a function of processing demands. *Journal of Experimental Child Psychology*, *61*, 242–275.

*Swanson, H. L., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence*, *21*, 83–108.

*Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities*, *39*(3), 252–269.

*Swanson, H. L., & Jerman, O. (2007). The influence of working memory on reading growth in subgroups of children with reading disabilities. *Journal of Experimental Child Psychology*, *96*, 249–283.

*Swanson, H. L., Saez, L., & Gerber, M. (2004). Do phonological and executive processes in English learners at risk for reading disabilities in Grade 1 predict performance in Grade 2? *Learning Disabilities Research and Practice*, *19*(4), 225–238.

Tannock, R. (1998). Attention deficit hyperactivity disorder: Advances in cognitive, neurobiological, and genetic research. *Journal of Child Psychology and Psychiatry*, *39*(1), 65–99.

*Van der Schoot, M., Licht, R., Horsley, T. M., Aarts, L. T., van Koert, B., & Sergeant, J. A. (2004). Inhibitory control during sentence reading in dyslexic children. *Child Neuropsychology*, *10*(3), 173–188.

*Van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology*, *87*, 239–266.

*Van der Sluis, S., van der Leij, A., & de Jong, P. F. (2005). Working memory in Dutch children with reading- and arithmetic-related LD. *Journal of Learning Disabilities*, *38*(3), 207–221.

Velluntino, F. R., & Fletcher, J. M. (2007). Developmental dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading; a handbook* (pp. 362–378). Oxford: Blackwell Publishing.

Velluntino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2–40.

Vukovic, R. K., & Siegel, L. (2006). The role of working memory in specific reading comprehension difficulties. In T. P. Alloway & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 89–112). Hove: Psychology Press.

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children- Fourth UK Edition (WISC-IV UK)*. London: Psychological Corporation.

*Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A., & Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, *110*(1), 157–172.

*Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology*, *27*(1), 35–78.

Appendix A: Number of studies located by differing search strategies

Search keywords	Search engine				
	Psych-info	Wilson web	Web of knowledge	Pub-med	Total
Executive function					
Children					
Reading	52	7	40	53	152
Reading difficulties	3	4	10	8	25
Dyslexia	6	4	11	25	46
Reading disability	19	6	31	9	65
Inhibition					
Children					
Reading	77	11	65	70	223
Reading difficulties	0	3	19	7	29
Dyslexia	16	2	21	23	62
Reading disability	19	4	39	12	74
Inhibitory skills					
Children					
Reading	0	1	8	6	15
Reading difficulties	0	0	2	2	4
Dyslexia	0	0	1	3	4
Reading disability	0	0	4	2	6
Working memory					
Children					
Reading	357	43	498	331	1,229
Reading difficulties	32	16	140	52	240
Dyslexia	47	10	165	183	405
Reading disability	94	10	144	42	290
Total	722	121	1,198	828	2,869

Note. Several papers were located using more than one search strategy and in more than one search engine.

Appendix B: Table indicating method of retrieval for studies included

Method of retrieval	Included
Psych-info	Altemeier *et al.* (2007), Brosnan *et al.* (2002), Jeffries and Everatt (2004), Nation *et al.* (1999), Reiter *et al.* (2005), Swanson and Ashbaker (2000), Swanson and Berninger (1995), Swanson *et al.* (2004), and Van der Sluis *et al.* (2004)
Pub-med	Cain (2004), Carretti *et al.* (2005), Censabella and Noel (2005), Helland and Asbjornsen (2000), Kupietz (1990), McGee *et al.* (1989), Miller-Shaul (2005), Van der Schoot *et al.* (2004), and Willcutt *et al.* (2001)
Web of knowledge	Bayliss *et al.* (2005), Cain and Oakhill (2006), Condor *et al.* (1995), De Berni and Palladino (2000), De Jong (1998), Howes *et al.* (1999), Kramer *et al.* (2000), Martinussen and Tannock (2006), Närhi and Ahonen (1995), Pennington *et al.* (1993), Roodenrys *et al.* (2001), Stothard and Hulme (1992), Swanson (1993), Swanson (1999), Swanson and Alexander (1997), Swanson and Jerman (2007), Swanson *et al.* (1996), and Willcutt *et al.* (2005)
Wilson web	McGee *et al.* (2004), Purvis and Tannock (2000), Savage and Frederickson (2006), Swanson *et al.* (2006), and Van der Sluis *et al.* (2005)
Reference citations	Everatt *et al.* (1997), Hall *et al.* (1997), Nyden *et al.* (1999), and Pickering and Gathercole (2004)
Author request	Everatt *et al.* (2008) and Protopapas *et al.* (2007)
Unpublished/in press/	Booth *et al.* (in press)

Note. Several papers were located in more than one search engine.

Appendix C: Details of tasks included in upper bound and lower bound analyses by modality and definitional criteria

	Criteria	Modality	Task	Study	Hedges' g (SE)	Z value
Upper bound	Discrepancy	Verbal	California Verbal Learning Test	Kramer et al. (2000)	0.18 (.16)	1.13
			Naming	Narhi and Ahonen (1995)	1.05 (.41)	2.56*
			Rapid automatic shifting	Altemeier et al. (2007)	0.96 (.18)	5.33***
			Sentence span	Willcutt et al. (2001)	0.36 (.18)	2.00*
		Non-verbal	Conflict task	Nyden et al. (1999)	0.87 (.47)	1.85
			Continuous performance task	Kupietz (1990)	1.05 (.45)	2.33*
				Pennington et al. (1993)	0.42 (.34)	1.24
			Five-point test	Reiter et al. (2005)	0.54 (.23)	2.35*
	Non-discrepancy	Verbal	Backward digit span task	Jeffries and Everatt (2004)	1.48 (.30)	4.93***
				Martinussen and Tannock (2006)	1.44 (.37)	3.89***
				McGee et al. (2004)	0.29 (.25)	1.16
				Savage and Frederickson (2006)	1.00 (.26)	3.85***
			Listening span	Cain and Oakhill (2006)	0.71 (.30)	2.37*
				Nation et al. (1999)	1.38 (.41)	3.37**
				Stothard and Hulme (1992)	0.55 (.39)	1.41
			Listening span – intrusions	De Berni and Palladino (2000)	0.92 (.43)	2.14*
			Object shifting task	Van der Sluis et al. (2004)	0.81 (.33)	2.46*
			Story recall	Swanson and Berninger (1995)	1.76 (.32)	5.50***
			Updating task	Carretti et al. (2005)	1.11 (.15)	7.40***
			Word recall	Cain (2004)	1.32 (.43)	3.07**
		Non-verbal	Coding task	Miller-Shaul (2005)	6.03 (.67)	9.00***
			Continuous performance task	Hall et al. (1997)	0.54 (.31)	1.74
			Numerical Stroop	Authors (under revision)	0.52 (.28)	1.86
			Stop task	Purvis and Tannock (2000)	0.41 (.35)	1.17
			Symbol search	Willcutt et al. (2005)	0.67 (.13)	5.15***
			Trail making task	McGee et al. (1989)	0.35 (.31)	1.13
			Visuospatial working memory task	Bayliss et al. (2005)	0.23 (.20)	1.15
	Unsure	Verbal	Auditory serial addition task	Roodenrys et al. (2001)	1.41 (.40)	3.53***
			Backward digit span	Pickering and Gathercole (2004)	0.32 (.19)	1.68

Q4

Q3

Heterogeneity of performance of children with reading difficulties on tasks of executive function 37

Appendix C: (*Continued*)

Criteria	Modality	Task	Study	Hedges' g (SE)	Z value
		Backward letter span	Howes et al. (1999)	0.83 (.26)	3.19**
		Letter naming	Swanson et al. (2006)	1.62 (.40)	4.05***
		Listening span	Van der Sluis et al. (2005)	0.19 (.31)	0.61
		Reading span	De Jong (1998)	1.62 (.38)	4.26***
		Rhyming task	Swanson and Jerman (2007)	1.36 (.35)	3.89***
			Swanson et al. (2004)	0.20 (.22)	0.91
		Semantic association	Swanson (1993)	0.85 (.26)	3.27**
		Sentence span	Swanson et al. (1996)	1.48 (.29)	5.10***
			Swanson (1999)	1.46 (.38)	3.84***
			Swanson and Ashbaker (2000)	2.68 (.36)	7.44***
		Sentential priming task (% errors)	Van der Schoot et al. (2004)	1.35 (.39)	3.46***
		Stroop	Everatt et al. (1997)	1.61 (.36)	4.43***
			Helland and Asbjornsen (2000)	1.20 (.30)	4.00***
			Protopapas et al. (2007)	0.86 (.30)	2.87**
		Verbal span	Everatt et al. (2008)	0.71 (.28)	2.54*
	Non-verbal	Concurrent digit colour	Swanson and Alexander (1997)	0.72 (.23)	3.13**
		Flanker letters	Censabella and Noel (2005)	1.57 (.42)	3.74***
		Group Embedded Figures Test	Brosnan et al. (2002)	1.15 (.28)	4.11***
		Tower of Hanoi	Condor et al. (1995)	0.72 (.34)	2.12*
Lower bound	Verbal	California Verbal Learning Test	Kramer et al. (2000)	0.18 (.16)	1.13
Discrepancy		Inhibition/switching task	Altemeier et al. (2007)	0.56 (.18)	3.11**
	Non-verbal	Continuous performance task	Kupietz (1990)	1.05 (.45)	2.33*
		Executive task	Närhi and Ahonen (1995)	0.45 (.39)	1.15
		Matching Family Figures Test	Pennington et al. (1993)	− 0.03 (.33)	− 0.09
		Stop-signal task	Willcutt et al. (2001)	0.13 (.15)	0.87
		Tower of London	Reiter et al. (2005)	− 0.13 (.22)	− 0.59
		Wisconsin Card Sorting Task	Nyden et al. (1999)	0.17 (.45)	0.38
Non-discrepancy	Verbal	Backward digit span	De Berni and Palladino (2000)	0.22 (.41)	0.54
		Delayed recall	McGee et al. (1989)	0.03 (.31)	0.10
		Digit reading task	Cain and Oakhill (2006)	0.15 (.30)	0.50

Appendix C: (*Continued*)

Criteria	Modality	Task	Study	Hedges' g (SE)	Z value
		Listening span	Stothard and Hulme (1992)	0.55 (.39)	1.41
		Object naming	Miller-Shaul (2005)	0.68 (.29)	2.35*
			Savage and Frederickson (2006)	0.19 (.24)	0.79
		Stroop	Jeffries and Everatt (2004)	0.04 (.27)	0.15
		Updating task (delayed intrusions)	Carretti et al. (2005)	1.11 (.15)	7.40***
		Verbal working memory task	Bayliss et al. (2005)	0.10 (.20)	0.50
		Word recall intrusion errors	Cain (2006)	1.32 (.43)	3.07**
	Non-verbal	Continuous performance task	Hall et al. (1997)	0.54 (.31)	1.74
			McGee et al. (2004)	0.05 (.24)	0.21
			Purvis and Tannock (2000)	−0.31 (.35)	−0.89
		Matrix	Swanson and Berninger (1995)	0.22 (.28)	0.79
		Reverse finger windows task	Martinussen and Tannock (2006)	1.07 (.34)	3.15**
		Spatial span	Nation et al. (1999)	0.01 (.37)	0.03
		Tower of London	Authors (*under revision*)	0.42 (.28)	1.50
		Trail making task	Van der Sluis et al. (2004)	0.08 (.32)	0.25
		Wisconsin Card Sorting test	Willcutt et al. (2005)	0.22 (.13)	1.69
Unsure	Verbal	Auditory digit sequence	Swanson and Ashbaker (2000)	0.01 (.26)	0.04
		Backward digit span	Van der Sluis et al. (2005)	0.07 (.31)	0.23
		Computation span	De Jong (1998)	0.45 (.34)	1.32
		Counting span	Swanson (1999)	1.20 (.36)	3.33**
		Letter generation	Swanson et al. (2004)	−0.17 (.22)	−0.77
		Listening span	Pickering and Gathercole (2004)	0.07 (.19)	0.37
		Listening sentence span	Swanson et al. (2006)	0.16 (.35)	0.46
		Number generation (2 s/item)	Roodenrys et al. (2001)	0.43 (.36)	1.19
		Phrase sequence	Swanson (1993)	0.22 (.25)	0.88
		Sentential priming task (% errors)	Van der Schoot et al. (2004)	1.35 (.39)	3.46**
		Stroop	Censabella and Noel (2005)	0.63 (.37)	1.70
			Everatt et al. (1997)	1.61 (.36)	4.43***
			Protopapas et al. (2007)	0.86 (.30)	2.87**

Q3

Appendix C: (*Continued*)

Criteria	Modality	Task	Study	Hedges' g (SE)	Z value
	Non-verbal	Updating task	Swanson and Jerman (2007)	0.64 (.32)	2.00*
		Facial memory	Howes et al. (1999)	0.24 (.25)	0.96
		Matrix	Swanson et al. (1996)	−0.28 (.28)	−1.00
			Swanson and Alexander (1997)	0.46 (.23)	2.00*
		Recognition	Brosnan et al. (2002)	0.49 (.36)	1.36
		Spatial memory	Everatt et al. (2008)	−0.32 (.28)	−1.14
		Tower of Hanoi	Condor et al. (1995)	0.72 (.34)	2.12*
		Wisconsin Card Sorting Task	Helland and Asbjornsen (2000)	0.41 (.27)	1.52

Appendix D: Descriptions of tasks identified

Task	Description
Abstract visual memory	Distinguish meaningless figures previously presented from six distractors. Non-verbal response
Animal test	Semantic fluency test – name as many animals as possible within 2 min
Arithmetic task	Subtest from WISC-r: solve a series of mental arithmetic problems (in addition to basic maths, this task provides a measure of verbal working memory)
Auditory digit sequence	Recall numbers previously presented in sentences in sequential order
Auditory serial addition task	Single digits presented every 2 s. Add each pair of successive numbers and immediately give the answer aloud
Backward digit span	Recall a set of digits in reverse order (lists of increasing length)
Backward letter span	Recall a set of letters in reverse order (lists of increasing length)
California Verbal Learning Test	Learn a 15 word list in 5 learning trials and complete a free recall test. An interference list is then presented and both free and cued recall tests given. Following a 20 min delay, there is a further free and cued recall test and also a recognition test. The score used was the number of false-positive errors, that is, the number of distractors incorrectly endorsed as this measures interference
Coding task	WISC III: transcribing a digit-symbol code as quickly as possible for 2 min
Computation span	Make a series of computations (either addition or subtraction) and after each computation a digit is presented. The presented digits had to be recalled in order
Concurrent digit colour	Cards with pictures of shapes were sorted into four piles. At the same time shown a different colour square every 2 s. Task is to point to the order of colour squares from an array of colours
Concurrent digit semantic	Digit sequences are presented. Sort cards into categories at the same time as listening to digits. Then asked to recall digits
Concurrent digit shape	Digit sequences are presented. Sort cards into piles placing identical pictures of shapes on top of each other at the same time as listening to digits. Then asked to recall digits
Conflict task	Responds twice when one stimuli is presented and once when two stimuli are presented
Contingency Naming Test	Three rows of nine different coloured stimuli. Each stimuli has large outside shape and a smaller inside shape (either congruent or incongruent). First – name either colour or large shape. Second – name colour if two shapes are the same or shape if two shapes are different. Third – rule the same but 9 stimuli have a backward arrow which indicates that the rule is reversed
Continuous performance task	Presented with 500 letters. First – press a key each time a white *S* is presented. Second – press the key only when a white *S* is followed by a blue *T*
Counting span	Count the number of dots presented in a series of arrays and recall the dot totals in serial order

Appendix D: (*Continued*)

Task	Description
Delayed recall – ROCF	Copy a complex figure using a different colour pencil for each section. There is a delay between recall trials
Digit naming	Name an array of digits as quickly as possible
Digit reading task	Read groups of three digits and recall the final digit
Executive task	Complete the WCST and the trail making task using the standard procedure. The Executive task score is based on the mean of *t* scores of WCST perseverative errors and the trail making task – part B
Facial memory	Recognize and identify black and white photos of faces of individuals of various ages, gender and ethnicity from a set of distractors
Five-point test	Connect the dots in a pattern of 5 symmetrically arranged dots with one or more straight lines to make as many different designs as possible
Flanker digits	Name a target digit flanked by either congruent or incongruent digits
Flanker letters	Name a target letter flanked by either congruent or incongruent letters
Flexibility task	A letter and a digit are presented at the same time on a computer screen. Press a button corresponding to the same side of the screen as the number and then press a button on the same side as the letter
Go/NoGo	Respond once every time two stimuli are presented
Group Embedded Figures Test	Locate a simple figure within a complex visual array ignoring distracters
Inhibition/switching test	Using the Stroop colour/word test: participants are required to switch between naming the colour in which words are printed and reading words that are printed within a box
Letter generation	Generate as many letters as possible in a non-systematic random order (requires inhibition of responses that would not be random)
Letter naming	Name an array of letters as quickly as possible
Listening span	Judge the veracity of sentences and then recall the last word from each sentence in sequence
Listening span – intrusions	Judge the veracity of sentences and then recall the final word from each sentence. Score is based on the number of intrusion errors
Mapping	Participant given a street map to study (lines and dots). Asked a process question and then asked to draw on another blank map, the lines and dots from the first map. Get progressively more complex
Matching Family Figures Test	Choose from six different pictures, one that is identical to a target picture
Matrix	Presented with a matrix with a number of shaded squares. Asked a processing question, then put an X in the squares which were shaded Given a matrix with dots in it to study. Asked a process question and then has to reproduce the pattern of dots
Memory updating (2 updates)	Repeat back the last three digits from a list

42   *Josephine N. Booth et al.*

Appendix D: (*Continued*)

Task	Description
Naming	First – name 50 letters presented in 5 rows. Second – name 50 letters, numbers and coloured squares presented in 5 rows in a random order
Non-verbal sequencing	Presented with a series of cards with pictures of shapes and line drawings. Organize the cards into rows; a certain number of cards have to be presented in each row. A process question is asked and the strategy used to remember the sequence is selected. Two distracter cards are inserted into the pack and the participant then has to reproduce the rows of cards
Number generation	Generate as many numbers as possible in a non-systematic random order
Number naming	Name each number presented in a table of 50 different numbers
Numerical Stroop – verbal	Name the number of x's or digits presented on a screen, the identity being either congruent or incongruent
Numerical Stroop-non-verbal	Underline specific numbers on a page and ignore distractors, inhibiting the first trial in the second trial
Object-inhibition-shifting task	Figure are presented with smaller figures inside. Participants had to alternate between naming the inner figure or outer figure depending on the colour
Object interference (Stroop)	Participants name blocks of colours on a page. Then they are presented with colour associated objects but printed in incongruent colours and instructed to name the colour of the ink
Object naming	Name an array of objects as quickly as possible
Object shifting	A series of geometric objects with a digit inside are presented. Participants have to either name the object or the figure depending on the colour of the stimuli
Object-inhibition task	A series of geometric objects with a smaller geometric object inside are presented. Participants have to inhibit the larger object and name the smaller object
Phrase sequence	An increasing number of phrases is presented, and the participant has to recall the phrase following a processing question
Picture sequence	A series of cards with pictures of shapes on them are presented. Following a process question, the participant has to arrange the cards in the correct order
Quantity inhibition task	Arrays of digits are presented which are incongruent to the actual digits (e.g. 222). Participants had to name the quantity in the array
Quantity naming task	Different numbers of triangles are presented and participants have to name the number (quantity) of shapes
Rapid automatic shifting	Alternate between rapidly naming a word and a double-digit number
Reading span	Read a series of sentences then store a presented word. Recall the presented words in order
Recognition task	A series of pictures is presented and a recall trial with distracters is carried out
Reverse finger windows task	Watch the examiner point to a series of windows on a card.

Q10

Q10

Appendix D: (*Continued*)

Task	Description
	Reproduce the sequence exactly, but in reverse order
Rhyming task	Listen to a set of words that rhyme. Then given a recognition task and asked to recall the previously presented words in order
Semantic association task	Organize sequences of words into abstract categories: presented with a set of words, then asked a discrimination question, then asked to recall the words which go together
Semantic categorization	One word presented every 2 s. Recall the category name for the list of words and then any word that went into that category
Sentence span	Participants provide the last word for a series of sentences and then have to recall the words
Sentential priming task	Judge whether the final word of a sentence is semantically congruent or incongruent – if congruent respond with the left hand, if incongruent respond with right hand. Had to inhibit response if pseudoword
Spatial memory	A series of black squares is presented. A specific series of squares is pointed to and the participant has to repeat the pointing. In the second condition, the participant has to point to the squares in reverse order
Spatial organization	Cards with varying shapes are presented. The participant has to replicate the correct series of cards
Spatial span	Rectangles on are presented on a screen. Each rectangle has three white squares with target stimuli in them. Participants have to indicate the odd one out by pointing. Stimuli moved across the screen and the participant has to recall the position of all of the odd one outs
Spatial working memory task	CANTAB – find hidden tokens while inhibiting responses to previous locations
Star Counting Test	Nine rows of stars presented with a number at the beginning of each row. Count the stars from top to bottom and left to right starting from this initial number. Plus and minus signs appear between some stars indicating subsequent stars should be counted either forward or backward from this point. In the first item, a plus sign indicates forward counting and a minus sign indicates backward counting, in the second item this is reversed
Stop signal	Letter X or O is presented on a computer monitor – press the corresponding key. In the second trial, inhibit response if a tone is presented
Story recall	Recall all the events in a story
Stroop	Read colour words printed in black ink, then name the colour of xxx's or blocks of colour, then name the colour of the ink in which an incongruent colour word is printed
S-word test	Name as many different words as possible beginning with the letter s
Symbol search	WISC III: deciding if target symbols appear in a row of symbols
Temporal order	A series of pictures is shown then followed with a delayed recall task

Appendix D: (*Continued*)

Task	Description
Tower of Hanoi	Three vertical posts and three different size disks. Move the pattern of disks to make different patterns following a series of rules
Tower of London	Three vertical posts of different heights and three different coloured balls (same size). Move the pattern of balls to make different patterns following a series of rules
Trail making task	Part a – use a pencil to connect a series of circles with numbers in them in ascending order. Part b – connect circles in ascending order alternating between numbers and letters (e.g. 1, A, 2, B, 3, C…)
Updating task	A series of one-digit numbers presented that varies in set lengths of 9, 7, 5, and 3. Recall the last three numbers presented
Updating task (delayed intrusions)	Pictures and nouns presented. Recall the highest (between positions 2–7) or lowest (between positions 9–14) pictures in the column that were named in the word list. Score is based on the number of delayed intrusion errors
Verbal fluency	Generate as many words as possible starting with a given letter (either s or f or a)
Verbal numerical Stroop	Given page of numbers, name every digit on page. Then in second part, have to say 'five' to the number 7 and vice versa
Verbal span	A series of digits of increasing lengths is presented. First – recall the digits in the same order; second – recall the digits in reverse order
Verbal working memory task	Nine different coloured squares forming a circle are presented with a digit in each square. Object names were presented and the participant has to think of the colour most associated with the object and touch the coloured square on the screen. Participants had to name the digit in the centre of the square they touched and recall all digits in serial order
Visuospatial working memory task	Nine different coloured squares forming a circle on it are presented with a digit in each square. Object names were presented and the participant has to think of the colour most associated with the object and touch the coloured square on the screen as quickly as possible. Participants had to remember the location of the squares they had touched and at the end recall all of the locations
Wisconsin Card Sorting Task	Sort cards either by colour, form, or number of shapes. Advised whether each sort is correct or not. After correct sorting of 10 cards the rule changes so that the sorting is based on another characteristic but participant not advised, must judge new sort by response of examiner
Word recall intrusion errors	Sets of concrete and abstract words are presented and recalled in correct serial order. The score is based on the number of intrusion errors

Author Queries

JOB NUMBER: 679
JOURNAL: BJDP

Q1   Please provide an alternative running head not more than 40 characters including spaces as per the journal style.

Q2   Please note that the reference citation Authors (in press) has been changed to Booth *et al.* (in press) with respect to the reference list provided.

Q3   References Roberts and Pennington (1996), Lezak (1995), Cain (2004), and Authors (under revision) have been cited in text but not provided in the list. Please supply reference details or delete the reference citations from the text.

Q4   Please suggest whether 'Z' and can be changed to 'z'.

Q5   Please check the edit of the sentence ' … it can by implied … ' to ' … it can be implied … '.

Q6   If possible, please supply DOI numbers for the journal articles in the list, as per current APA style.

Q7   Please check the inserted author names in reference Booth *et al.* (in press).

Q8   Please update publication details in reference Booth *et al.* (in press).

Q9   Reference Burgess *et al.* (2006) is provided in the list but not cited in the text. Please supply citation details or delete the reference from the reference list.

Q10   Please suggest whether 'distracter' can be changed to 'distractor' for consistency.