THE JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY
ACAMH
Journal of Child Psychology and Psychiatry 54:2 (2013), pp 131–143 doi:10.1111/jcpp.12001

Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct?

Maggie E. Toplak,[1] Richard F. West,[2] and Keith E. Stanovich[3]

[1]Department of Psychology, LaMarsh Centre for Child and Youth Research, York University, Toronto, ONT, Canada;[2]Department of Graduate Psychology, James Madison University, Harrisonburg, VA, USA;[3]Human Development and Applied Psychology, University of Toronto, ONT, Canada

Background: Both performance-based and rating measures are commonly used to index executive function in clinical and neuropsychological assessments. They are intended to index the same broad underlying mental construct of executive function. The association between these two types of measures was investigated in the current article. **Method and Results:** We examined the association between performance-based and rating measures of executive function in 20 studies. These studies included 13 child and 7 adult samples, which were derived from 7 clinical, 2 nonclinical, and 11 combined clinical and nonclinical samples. Only 68 (24%) of the 286 relevant correlations reported in these studies were statistically significant, and the overall median correlation was only .19. **Conclusions:** It was concluded that performance-based and rating measures of executive function assess different underlying mental constructs. We discuss how these two types of measures appear to capture different levels of cognition, namely, the efficiency of cognitive abilities and success in goal pursuit. Clinical implications of using performance-based and rating measures of executive function are discussed, including the use of these measures in assessing ADHD. **Keywords:** ADHD, executive function assessment, performance-based measures, ratings of behavior.

Introduction

Executive function is one of the most widely invoked constructs in the cognitive science, neuropsychology, developmental, and clinical research literatures. The operationalization and measurement of executive functions is a key issue that directly impacts the inferences we can make about these competencies. The procedures used to operationalize executive function in clinical settings employ either performance-based or rating measures. Performance-based measures involve standardized procedures that are administered by an examiner and usually assess accuracy and/or response time. Rating measures of executive function involve an informant reporting on difficulties with carrying out everyday tasks. It currently remains unclear to what extent performance-based and ratings of executive function assess the same underlying construct. The purpose of this study was to examine the relationship between these two types of measures of executive function. The goal of this study is to provide clinicians with a perspective informed by theory and research for the use of these measures in the context of a clinical assessment.

The relevance of executive functions in clinical assessment

The construct of executive functions has become important in the assessment of typically developing children (Davidson, Amso, Anderson, & Diamond, 2006), aging adults (Salthouse, Atkinson, & Berish, 2003) and special populations (Nigg, 2006; Pennington, 2002). Executive functions are assumed to play an important role in the efficiency of goal-directed behavior (Miyake, Friedman, Emerson, & Witzki, 2000; Pennington & Ozonoff, 1996; Salthouse et al., 2003; Strauss, Sherman, & Spreen, 2006), especially, in novel contexts where there are no well-learned behaviors to draw upon (Shallice, 1990). Several elements of executive function have been articulated as key components, including: anticipation and deployment of attention, impulse control and self-regulation, initiation of activity, working memory, mental flexibility and utilization of feedback, planning ability and organization, and selection of efficient problem-solving strategies (Anderson, 2008). However, the most typical domains assessed as indices of executive function are updating (constant monitoring and rapid addition/deletion of working memory contents), shifting (switching flexibly between tasks or mental sets), and inhibition (deliberate overriding of dominant or prepotent responses; Miyake & Friedman, 2012). A major challenge in the assessment of executive functions is the impurity problem, namely that most measures of executive function involve non-executive processes in the task context, such as color naming in the Stroop task (Miyake & Friedman, 2012).

Executive processes develop and change over the lifespan (Davidson et al., 2006; Lamm, Zelazo, & Lewis, 2006; Williams, Ponesse, Schachar, Logan, & Tannock, 1999), but individual differences in execu-

Conflicts of interest statement: No conflicts declared.

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA 02148, USA

tive functions show relative stability over the course of development (Miyake & Friedman, 2012). Deficits in executive functions are of particular relevance for clinicians. Such deficits may result in inappropriate social behavior, problems with decision-making and judgment, and difficulties with initiating, following, shifting, and organizing plans (Damasio, 1994, 1996; Strauss et al., 2006). Difficulties with executive functions have been implicated in several neurological and psychiatric conditions such as: traumatic brain injuries (Clark, Manes, Antoun, Sahakian, & Robbins, 2003; Labudda et al., 2009); schizophrenia (Cavallaro et al., 2003; Kester et al., 2006; Nakamura et al., 2008); substance use (Barry & Petry, 2008; Ernst et al., 2003); obsessive-compulsive disorder (Lawrence et al., 2006); psychopathy (Mahmut, Homewood, & Stevenson, 2008); attention-deficit/hyperactivity disorder (Toplak, Jain, & Tannock, 2005); and pathological gambling (Toplak, Liu, MacPherson, Toneatto, & Stanovich, 2007).

Both performance-based and rating measures of executive function have been used to clinically assess many of the conditions listed above. However, the extent to which these two types of measures actually reflect the same underlying mental construct is far from certain and has never before been examined in a review. In the analysis, we report here, we draw heavily on the literature on ADHD, where a substantial number of studies have used both performance-based and ratings measures of executive function.

ADHD is a neurodevelopmental disorder that has been characterized by deficits in executive function (Barkley, 2006; Nigg, 2006). Performance-based measures of executive function have found reliable decrements in the performance of ADHD as compared with control groups (Barkley, 2006; Nigg, 2006; Nigg et al., 2005; Scheres et al., 2004; Sergeant, Geurts, & Oosterlaan, 2002). Rating measures of executive function have also found reliable differences, as individuals with ADHD are typically rated as having more difficulties with everyday tasks presumed to involve executive function processes (Barkley & Fischer, 2011; Barkley & Murphy, 2010a,b, Barkley and Murphy 2011; Biederman et al., 2008; Hummer et al., 2010; Mahone et al., 2002; Toplak, Bucciarelli, Jain, & Tannock, 2008). Thus, both performance-based and rating measures of executive function have been found to reliably differentiate between ADHD and control groups. However, relatively little attention in the literature has focused on explaining the relationship between the performance-based and rating measures themselves. Specifically, the question of whether these measures index the same or different underlying mental construct remains largely unexamined.

The purpose of this study was to assess the relationship, or lack thereof, between performance-based and rating measures of executive function. We begin this analysis by considering the administration characteristics of performance-based and rating measures of executive function. The quantitative relationship between these measures is then examined. That is, the associations between performance-based and rating measures of executive function are examined in studies with clinical and nonclinical samples. Finally, theoretical perspectives from both clinical research and cognitive science are considered. As performance-based and rating measures of executive function have been frequently examined in the ADHD literature, the clinical implications of this relationship (or lack of relationship) are considered.

Performance-based measures versus ratings of executive function

Performance-based measures of executive function

The conventional measurement of executive function has been based on cognitive performance-based tests (Pennington & Ozonoff, 1996). Performance-based tests are administered in highly standardized conditions. Stimulus presentation is carefully controlled so that each examinee experiences and completes the task in precisely the same way as other examinees. In addition, the measures of performance are typically based on the examinee's accuracy, response time, and/or speeded responding under a time constraint. There are several performance-based measures of executive function, such as: the Wisconsin-Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtis, 1993), the Stroop test (Jensen & Rohwer, 1966; MacLeod, 1991; Stroop, 1935), and tests of verbal fluency (Strauss et al., 2006). The WCST requires the maintenance of a task set, flexibility in response to feedback, avoiding perseverative tendencies, and inhibiting a prior response that is no longer appropriate (Salthouse et al., 2003). The Stroop effect (MacLeod, 1991; Stroop, 1935) is a demonstration of interference control. In the Stroop test's key condition, the participant must inhibit an overlearned response (reading a word that names a color) to respond with another dimension that is incongruent and 'interfering' (naming the ink color of the word, instead of the actual color word). Verbal fluency tests require the maintenance of a task set (generating items that fit a particular criteria or category), generating multiple responses, monitoring and avoiding repetitions, and using different retrieval strategies (Salthouse et al., 2003).

Although this is only a sample of available performance-based tests of executive function, these tests share the same general characteristics (see Strauss et al., 2006 for a comprehensive list of performance-based measures of executive function). They are all administered under highly standardized conditions with a single examiner who provides specific feedback or direct prompts to the examinee to direct performance. Accuracy and response time are the typical dependent measures on these tests. A

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

key dependent measure on the WCST is the total number of sets of 10 consecutive correct pairings. On the Stroop, the typical measure is the difference between the response time for naming the ink colors in the incongruent condition minus the response time for naming the actual ink colors. The key dependent measure on the verbal fluency test is the total number of items given by the examinee in the period of one minute. These measures and several others have been examined in studies in conjunction with ratings of executive function; all these studies have been included in this review.

Rating scales of executive function

Rating scales of executive function were developed to provide an ecologically valid indicator of competence in complex, everyday, problem-solving situations (Roth, Isquith, & Gioia, 2005). An assumption underlying the use of these rating scales is that they are measuring behaviors that are importantly related with processes that are assessed by performance-based measures of executive function. Our literature review found that the most commonly used rating scale of executive function has been the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). This instrument is composed of eight individual scales and three composite scores. The Inhibit, Shift, and Emotional Control scales compose the Behavioral Regulation Index Composite. The Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor scales compose the Metacognition Index Composite. The Behavior Regulation and Metacognition Indices can be combined to form an overall Global Executive Composite. Examinees respond to a total of 86 items that describe difficulties in everyday activities. For example, some of the items indicate situations in which impulses were not controlled (Inhibit scale), items describing difficulty with staying with an activity (Working Memory scale), or items describing difficulties encountered with a disorderly workspace (Organization of Materials scale). Each item is rated on whether difficulties are encountered: Never, Sometimes, or Often. A score for each Index and Composite can be used to derive a scaled score that indicates degree of difficulty in each rated domain. The only constraint given to informants is to report on behaviors that have been problematic in the last 6 months. There are two validity scales to assess for Inconsistency and Negativity of ratings. There are also preschool and adult versions of this particular scale (Gioia, Espy, & Isquith, 2003; Roth et al., 2005).

One characteristic of the BRIEF is that some of its scales have no parallel performance-based measures of executive function. Although the Inhibit and Working Memory scales map onto parallel performance-based measures of executive function, the Initiate and Organization of Materials scales do not map onto analogous performance-based measures. Other rating scales of executive function more closely mirror performance-based measures, such as the Childhood Executive Functioning Inventory (CHEXI), which has only inhibition and working memory scales (Thorell, Eninger, Brocki, & Bohlin, 2010; Thorell & Nyberg, 2008).

Table 1 provides a list of several behavior rating scales of executive function. Some of these measures are proprietary tests; whereas others are rating scales described in detail in peer reviewed articles or books. Most of these measures have been developed in the context of clinical settings. A comprehensive review of ratings of executive function can be found in Malloy and Grace (2005).

The measurable association between performance-based and ratings of executive function

If performance-based and rating measures of executive function are assessing the same general construct, then these measures should be strongly positively correlated. That is, high competence measured by ratings should be associated with high competence on performance-based measures. To examine this, we conducted *PsyInfo* and *Pubmed* searches that identified 20 empirical studies that tested the association between performance-based and rating measures of executive function based. Our approach was to be as inclusive as possible, so that the analysis would include a range of studies with different periods of development (child and adult) and different populations (clinical and nonclinical samples). Thirteen of the 20 studies were conducted with children and seven were conducted with adults. Seven studies reported results based on clinical samples, 11 reported results based on combined clinical and nonclinical samples, and two studies reported results based on nonclinical adult samples. Sixteen studies reported correlational analyses and four only reported that uniformly nonsignificant correlations were found, without reporting actual r or p values. Each of the 20 studies is marked with an asterisk in the reference section. Details of the sample, measures, and results for each study are reported in online Table S1.

In the studies that reported correlational analyses, the following rating scales of executive function were used: 13 used the BRIEF, five used the Behavioral Assessment of the Dysexecutive Syndrome-Dysexecutive Questionnaire (BADS-DEX), and three used an impulsivity scale as an index of a lack of inhibition (one study used both the BRIEF and BADS-DEX). Numerous performance-based measures that tapped a variety of different aspects of executive function were used in these studies: working memory (verbal and nonverbal), planning, mental flexibility, perceptual/motor planning, response inhibition, resistance to distraction, and set-shift-

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

Table 1 List of rating measures of executive function and frontal processes

Measure	Ages	Subscales
Behavior Rating Inventory of Executive Function (BRIEF) – Parent, teacher, and self-report forms (Gioia et al., 2000, 2003; Guy et al., 2005; Roth et al., 2005)	Preschool version: parent form Child/adolescent version: parent and teacher forms Self-report version Adult version	Inhibit, shift, and Emotional Control Scales form the Behavioral Regulation Index Initiate, working memory, plan/organize, Organization of Materials, and Monitor Scales form the Metacognition Index Behavioral Regulation Index and Metacognition Index form a Global Executive Composite
Brown Attention-Deficit Disorder Scales for adolescents and adults (Brown, 2001)	Adolescent and adult forms	Activation, focus, effort, emotion, memory, and action subscales. Composite score also available
Childhood Executive Functioning Inventory (CHEXI; Thorell & Nyberg, 2008; Thorell et al., 2010)	Parent and teacher reports for children	Working memory and inhibition subscales.
Current Behavior Scale (CBS; Barkley, 1997; (items available in Biederman et al., 2008; original source unpublished)	Self report for adults	Total score for executive function deficits
Deficits in Executive Function Scale (Barkley & Murphy, 2010a,b)	Self report for adults and other report	Five scales: self-management to time, self-organization and problem-solving, self-discipline, self-motivation, and self-activation
Dysexecutive Questionnaire (DEX); Part of the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al., 1996) for adults and BADS-C (Emslie et al., 2003)	Adult version has self-report and other respondent forms; Child version completed by parent and teacher	Single Scale designed to measure emotional/personality, motivational, behavioral, and cognitive changes
Executive Function Index (Miley & Spinella, 2006; Spinella, 2005)	Adult self-report scale	Five scales: empathy, strategic planning, organization, impulse control, and motivational drive
Frontal Behavior Inventory (FBI; Kertesz et al., 1997)	Adult scale. Structured interview by clinician with patient's caregiver as informant	Single Scale
Frontal Systems Behavior Scale (FrSBe; Grace & Malloy, 2001)	Adult. Self-rating and family rating form	Questionnaire intended to measure adult behavior before and after frontal systems damage. There are three subscale scores of apathy, disinhibition, and executive dysfunction, and a total score
Iowa Rating Scales of Personality Change (IRSPC; Barrash et al., 2000)	Adult self-report	Single Scale
Neuropsychiatric Inventory (NPI; Cummings et al., 1994)	Adult scale based on caregiver interview	Single Scale
Working Memory Rating Scale (WMRS; Alloway et al., 2008)	Teacher report	Single composite measure of working memory deficits

ing/mental flexibility. Within each of the performance-based constructs, different measures were often used to assess the same construct. For example, the Digit Span subtest and the N-back task were both used as indicators of working memory. We decided that the most coherent way to organize this varied set of studies was to group them on the basis of the rating scale of executive function used. A summary of correlational analyses from each study is included in Table 2.

Behavior Rating of Executive Function (BRIEF)

Thirteen studies examined the association between the BRIEF questionnaire and performance-based measures of executive function. Eight of these studies included children (Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Brown et al., 2008; Conklin, Salorio, & Slomine, 2008; Mahone et al., 2002; Mangeot, Armstrong, Colvin, Yeates, & Taylor, 2002; McAuley, Chen, Goos, Schachar, & Crosbie, 2010; Parrish et al., 2007; Vriezen & Pigott, 2002), three included adolescents (Hummer et al., 2010; Niendam, Horwitz, Bearden, & Cannon, 2007; Toplak et al., 2008), one included both children and adolescents (Bodnar, Prahme, Cutting, Denckla, & Mahone, 2007), and one included young adults (Shuster & Toplak, 2009). These studies included clinical samples (five studies), a clinical sample and control group (seven studies), or a nonclinical sample (one study). Clinical samples included medical or neurological conditions (phenylketonuria, hydrocephalus, spina bifida, traumatic brain injury, orthopedic injury, or epilepsy) or psychiatric conditions (psychosis risk, ADHD, or Tourette's Syndrome).

There were multiple dependent measures and multiple comparisons within a given study. The 13

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

doi:10.1111/jcpp.12001

Table 2 Summary of correlations between performance-based and rating measures of executive function reported within clinical, control group, and total (clinical and control groups collapsed) samples

Study	Performance-based executive function measure	Clinical group	Control group	Total group
Studies using the BRIEF Scale				
Anderson et al. (2002)	Planning: Tower of London Mental flexibility: contingency naming test Verbal fluency: controlled oral word association test Perceptual planning: Rey complex figure	NR	NR	From .01 to .48***
Bodnar et al. (2007)	Inhibitory control and sustained attention: Conners' continuous performance test II and test of variables of attention	From .01 to .31**	–	–
Brown et al. (2008)	Executive function/attention: Conner's continuous performance test and children's category test	NR	NR	From ns to .23*
Conklin et al. (2008)	Working memory: Digit Span Backwards Age Scaled Score	.17	–	–
Hummer et al. (2010)	Inhibition: Stroop test, counting interference test Working memory: Children's Memory Scale Numbers subtest Inhibition and variability in attention: Conners' continuous performance test	NR	NR	From ns to –.37**
Mahone et al. (2002)	Fluency: Controlled oral word association test Planning: Tower of London Inhibition and Persistence: Tests of Variables of Attention – Visual (TOVA)	NR	NR	From .17 to .36***
Mangeot et al. (2002)	Executive function measures: Consonant trigrams, contingency naming test, word fluency test, underlining test, Rey-Osterreith complex figure	From .03 to .26*[a]	–	–
McAuley et al. (2010)	Inhibition and performance monitoring: stop signal task Working memory: N-back task	NR	NR	From .01 to .26*
Niendam et al. (2007)	Executive function tests: Matrix reasoning, Trailmaking Test part B, verbal letter fluency	ns	–	–
Parrish et al. (2007)	Executive Function Delis–Kaplan Executive Function System (D–KEFS) subtests: sorting test, verbal fluency test, and color-word interference	NR	NR	From .11 to .33*
Shuster and Toplak (2009)	Inhibition: stop task and Stroop test	–	–	From –.32* to .09
Toplak et al. (2008)	Inhibition: stop task Working memory: Digit Span and Spatial Span composites Set-shifting: trailmaking part B time Planning: Stockings of Cambridge	NR	NR	From .09 to .41***
Vriezen and Pigott (2002)	Executive function: Wisconsin card sorting test, Trailmaking Part B time, verbal fluency tests	From .03 to .26***	–	–
Studies using the BADS DEX Questionnaire				
Bennett et al. (2005)	BADS 6 subtests: rule shift card test, action program test, key search test, temporal judgment test, zoo map test, and modified six elements test	From .03 to .51**	–	–
Burgess et al. (1998)	Executive function tests: modified Wisconsin card sorting test, cognitive estimates test, verbal fluency, Trailmaking Time for parts A and B, simplified six element test	From .00 to .40***	NR	NR
Norris and Tate (2000)	Executive function tests: Wisconsin card sorting test, Trailmaking Test part B, Rey-Osterreith complex figure test, cognitive estimation test, controlled word association test, BADS battery of six subtests	From .06 to .41*	NR	NR
Odhuba et al. (2005)	Response initiation and suppression: Hayling test Rule detection and response flexibility: Brixton test	From .15 to .48*	–	–
Wilson et al. (1998)	Executive function tests: BADS battery of six subtests	NR	NR	NR to .62***

Table 2 *Continued*

Study	Performance-based executive function measure	Clinical group	Control group	Total group
Studies using impulsivity rating (lack of inhibition)				
Enticott et al. (2006)	Inhibition measures: inhibitory reach task, stop signal task, spatial Stroop task, negative priming task	–	From −.17 to .56***	–
Riccio et al. (1994)	Cognitive flexibility: Wisconsin card sorting test	From .21 to .30	–	–
Shuster and Toplak (2009)	Inhibition: stop task and Stroop test	–	From .05 to .07	–

NR, not reported; ns, not significant.
Correlations have been computed so that good performance on each measure is indicated by a positive score.
[a]Reported as standardized regression co-efficient in Mangeot et al. (2002).
*$p < .05$, **$p < .01$, ***$p < .001$.

studies produced a total of 306 possible correlations, of which 182 actual correlations were reported for examination. Of these 182 reported correlations, only 35 were statistically significant (19%). For those studies that reported *r*-values, the mean correlation was .15 and the median reported correlation coefficient was .18. These values are likely to provide an overestimation of the values that would have been obtained had all 306 possible correlations been reported, because a number of the studies did not report values for nonsignificant correlations. On the basis of these studies, the association between ratings on the BRIEF and performance-based measures of executive function seems to be extremely weak.

Behavioral Assessment of the Dysexecutive Syndrome-Dysexecutive Questionnaire (BADS-DEX)

The BADS has two major components: A battery of performance-based measures of executive function and the 20-item DEX that yields a single score from each informant. Five studies including adult-only samples examined the association between the BADS-DEX questionnaire and performance-based measures of executive function (Bennett, Ong, & Ponsford, 2005; Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Norris & Tate, 2000; Odhuba, van den Broek, & Johns, 2005; Wilson, Evans, Emslie, Alderman, & Burgess, 1998). Three of these studies contained clinical sample and control groups, and two of them contained only a clinical group. The clinical samples included medical or neurological conditions (rehabilitation patients, traumatic brain injury, multiple sclerosis, or other neurological disorders) or psychiatric conditions (schizophrenia).

The five studies reported all of the 76 possible relevant correlations that resulted from the various multiple dependent measures and multiple comparisons. Only 28 (37%) of these correlations were statistically significant. The 76 reported correlations had an overall mean value of $r = .14$ and median value of $r = .14$. Thus, only a very weak association was found between ratings on the BADS-DEX questionnaire and performance-based measures of executive function.

Impulsivity as a measure of a lack of inhibition

Three of the studies examined the association between impulsivity ratings (as indicators of a lack of inhibition) and performance-based measures of inhibition. Two studies were conducted with an adult nonclinical sample (Enticott, Ogloff, & Bradshaw, 2006; Shuster & Toplak, 2009) and one study was conducted with a child sample referred for behavioral and learning problems (Riccio, Hall, Morgan, Hynd, & Gonzalez, 1994). Overall, 28 of the 34 possible relevant correlations were reported in these studies. Five (19%) correlations were statistically significant. The mean was $r = .21$ and median was

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

r = .25. These findings suggest that ratings of impulsivity are very modestly related to performance-based measures of executive function.

Summary of the measurable associations between performance-based and rating measures of executive function

We scrutinized 20 studies that examined the association between performance-based and ratings measures of executive function. Amalgamated across studies, only 68 (24%) of a total of 286 correlational comparisons were statistically significant. The magnitude of correlations obtained was quite low, with median values of r = .18, .14, and .25 for the BRIEF, BADS-DEX questionnaire, and the impulsivity rating measures, respectively. The median correlation was only 0.19 across all these studies. Even these values are likely to be on the high side, because some of the studies did not report values for nonsignificant correlations. Given that the great majority of the correlations reviewed in this study were not significant, it is likely that Type 1 error was responsible for some the correlations reaching a level of significance. Furthermore, via the well-known file-drawer problem (that significant effects are differentially advantaged in the publication process), there may have been a number of nonsignificant findings that were never published. The results of the studies that we analyzed revealed a surprising lack of association between performance-based and ratings of executive function. The small to modest association is unlikely to be the result of self- versus other-ratings or to clinical versus nonclinical samples, because the pattern of associations was comparable across these studies and samples. Although both types of measures are supposed to index the same underlying mental construct, a basic principle of convergent validity in science is that different operational measures of the same construct should correlate highly. This is apparently not the case for performance-based and rating measures of executive function.

Theoretical perspectives on what performance-based and ratings of executive function measure

From the perspective of operationalization, performance-based and rating measures of executive function are different in terms of how they are administered and scored. Our review of the existing empirical literature indicates that the two types of measures are also only minimally correlated. We suggest that performance-based and ratings of executive function assess different aspects of cognitive and behavioral functioning that independently contribute to clinical problems.

We might begin to explain the lack of association between performance and ratings of executive function by drawing an analogy with the field of intelligence. Like the case of executive functioning, the construct of intelligence has also been defined broadly, but measured narrowly. This distinction between broad and narrow theories of intelligence is discussed by Stanovich (2009b), who noted that 'broad theories include aspects of functioning that are captured by the vernacular term intelligence (adaptation to the environment, showing wisdom and creativity, etc.), whether or not these aspects are actually measured by existing tests of intelligence. Narrow theories, in contrast, confine the concept of intelligence to the set of mental abilities actually tested on extant IQ tests' (p. 12). That is, a full-scale intelligence score (narrow sense) does not assess all of the ways that someone may be considered to be 'smart' as a layperson might understand that term (broad sense). This is analogous to executive functions. For example, few perseverative errors on the WCST (narrow sense) does not index all the ways that someone shows competence in novel problem-solving and goal-directed behavior (broad sense).

Our explanation for the lack of convergence displayed by the performance-based and behavior rating measures involves positing that these measures are actually tapping different cognitive levels—specifically, what Stanovich (2009b, 2011) terms the difference between the algorithmic and the reflective mind. Cognitive scientists refer to the level of analysis concerned with efficiency as the algorithmic level of analysis (Anderson, 1990; Marr, 1982; Stanovich, 1999, 2009b). The cognitive psychologist and neuropsychologist work largely at this level by showing that human performance can be explained by information processing mechanisms in the brain, such as, input coding mechanisms, perceptual registration mechanisms, working memory, long-term memory, etc. In contrast, the reflective level of analysis is concerned with the goals of the person, beliefs relevant to those goals, and the choice of action that is rational given the goals and beliefs (Bratman, Israel, & Pollack, 1991; Dennett, 1987; Newell, 1982, 1990; Pollock, 1995; Stanovich, 2009b, 2011). In short, the reflective level is concerned with the goals of the system, beliefs relevant to those goals, and the choice of action that is optimal given the system's goals and beliefs. It is only at the level of the reflective mind where issues of optimal decision-making come into play.

The important distinction between the algorithmic level of analysis and the reflective level from cognitive science maps analogously onto importantly differentiating performance-based from rating measures of executive function. Only the latter assess issues of rational control, which refers to behavior in the real environment that serves to foster goal achievement. Performance measures may indeed be assessing something of genuine importance, namely the efficiency of the processes available to recruit in behavioral control, such as inhibition, but performance-based measures bypass the whole issue of

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

rational goal pursuit. This point about the laboratory measures has been made before by Salthouse et al. (2003): 'The role of executive functioning may also be rather limited in many laboratory tasks because much of the organization or structure of the tasks is provided by the experimenter and does not need to be discovered or created by the research participant' (p. 569). Performance-based measures of executive function provide important information regarding efficiency of processing, but ratings of executive function tell us more about success in rational goal pursuit.

It is extremely important to differentiate between the algorithmic and reflective levels, as they provide different information about cognitive functioning. For this reason, Stanovich (2009a) has suggested that the term executive processes has been misnamed. The term 'executive' conflates these two different levels and 'mistakenly implies that everything 'higher up' has been taken care of, or that there is no level higher than what these executive functioning tasks measure' (Stanovich, 2009a, p. 67). Performance-based tasks would be better described as supervisory processes, as regulation is directed by an external examiner.

The conceptual differentiation of performance-based and rating measures of executive function is also consistent with an important distinction in psychometrics. Psychometricians have long distinguished typical performance situations from optimal or maximal performance situations (see Ackerman, 1994, 1996; Ackerman & Heggestad, 1997; Cronbach, 1949; Matthews, Zeidner, & Roberts, 2002). Typical performance situations remain unconstrained in that no overt instructions to maximize performance are given, and the task interpretation is determined to some extent by the participant. The goals to be pursued in the task are left somewhat open. The issue is what a person would typically do in such a situation, given few constraints. Typical performance measures assess in part goal prioritization and epistemic regulation. In contrast, optimal performance situations are those where the task interpretation is highly constrained, and the person performing the task is instructed to maximize performance. Thus, optimal performance measures examine questions of the *efficiency* of goal pursuit. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of critical thinking and cognitive styles are often assessed under typical performance conditions. Likewise, many measures of rational thinking and decision-making are assessed under typical performance conditions (Stanovich, 2009b; Stanovich, West, & Toplak, 2011).

Performance-based and ratings measures of executive function cleave the optimal/typical distinction in different ways. It is clear that performance-based measures are assessed under optimal/maximal conditions. This characteristic of neuropsychological tests of executive function has been echoed by Gioia, Isquith, and Kenealy (2008), who argue that 'individuals with substantial executive dysfunction can often perform adequately on well-structured tests when the examiner is allowed to cue and probe for more information, relieving the individual of the need to be appropriately inhibited, flexible, strategic in planning, and goal directed.' (p. 180). However, both performance-based measures of executive function capture optimal performance situations, because the task interpretation is determined externally by the examiner and is not left up to the participant.[1]

In contrast, ratings of executive function are unlike measures of maximal or optimal performance. When participants are estimating the frequency and typicality of how well they perform in day-to-day situations that are likely to engage executive processes, their responses are not constrained by an external examiner and there are no explicit instructions to 'maximize' or 'optimize' their ratings. The interpretation of the task is left up to the rater, who must decide on instances from their everyday lives that map onto the questions or constructs probed. Their task is to provide an estimate of the frequency of such events. Ratings of behaviors related to executive functions are also fraught with challenges related to informant reports, such as context effects and differences in the way different observers judge behavior (Barkley, 2006). Both performance-based and rating measures of executive function provide important and nonredundant information about an individual's efficiency and success in achieving goals.

Implications for the use of performance-based and rating scale measures of executive function in ADHD

The findings from our review of the empirical studies that have examined the association between these measures indicate that the two different types of measures are assessing different aspects of cognitive functioning. Studies from the field of ADHD are consistent with this conclusion. Biederman et al. (2008) examined overlap in impairment on perfor-

[1] It is important to note that all measures of executive function may not cleanly 'cleave' the optimal/typical distinction, as described in this study. For example, the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al., 1998) was included as a performance measure administered under maximal conditions, because it is administered under highly structured conditions by an examiner. However, this measure was developed to provide a more ecologically valid clinical tool for assessing executive functions. For this reason, some of its subtests may be more like tests of typical performance. For example, the Temporal Judgement Test asks the examinee to estimate how long certain activities take in real life. Performance on this particular subtest has been found to be unrelated to other conventional performance-based tests of executive function (Norris & Tate, 2000).

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

mance-based and ratings measures of executive function in a sample of adults with ADHD. There was little overlap in the impairment identified by these two domains of measures. Only 14% of the participants who had impairment on the performance-based measures of executive function reported impairment on the rating scale measure of executive function. Also, Barkley and Murphy (2010a,b) explored whether performance-based and ratings measures of executive function made overlapping or unique contributions to explaining self-reported occupational problems in a sample of adults with ADHD. In fact, the two different types of measures of executive function explained separate variance in occupational success in adults with ADHD.[2]

The distinction between typical and maximal performance maps onto other theoretical models of ADHD that distinguish between motivational and executive processes (Sonuga-Barke, 2002). The distinction between these domains have been shown behaviorally (Crone, Jennings, & van der Molen, 2003; Martel & Nigg, 2006; Shuster & Toplak, 2009; Sonuga-Barke, Dalen, & Remington, 2003) and has also been articulated in neural models (Sagvolden, Johansen, Aase, & Russell, 2005). As Barkley (1997) has noted, 'measures taken in clinics or laboratory assessments over relatively brief temporal durations are going to prove less sensitive to the identification of the disorder and its associated cognitive deficits than will measures collected repeatedly over longer periods of time…' (p. 332). The impairments of ADHD manifest as typical performance in day-to-day activities, across different situations and contexts (which is part of the diagnostic criteria in DSM-IV-TR, American Psychiatric Association, 2000). An implicit assumption in ratings of typical day-to-day behavior is that the subject's natural tendencies to internally regulate his/her behavior are being assessed—namely how well he/she can carry out these activities without constant direction or regulation by an external evaluator. It is apparent from these studies that both performance-based and ratings of executive function provide important but distinctive types of information with respect to ADHD behavior. Specifically, performance-based measures assess the processing efficiency of cognitive abilities, whereas ratings of executive function assess the extent to which the individual is achieving his/her goals. They assess different aspects of functioning.

The implication for the practicing clinician is that there is more separability than commonality among performance-based and ratings of executive function. These two classes of measures should not be interpreted as equivalent, interchangeable, or as types or subcategories of one another. The fact that both sets of measures are defined as executive functions in name further confuses the issue, suggesting that these measures are alike, when in fact they represent different aspects of cognitive and behavioral functioning. Impairment on performance-based measures of executive function does not translate into impairment on ratings of executive function, or *vice versa*, as was demonstrated by Biederman et al. (2008). It remains an empirical question to determine the correlates, convergers, and predictive utility of both performance-based and rating measures of executive function. Such research will be useful to properly characterize these measures. For example, there is already some literature indicating that ratings of executive function are significantly related to impairment in major life activities and in occupational functioning in adults (Barkley & Fischer, 2011; Barkley & Murphy, 2010a,b, 2011).

Performance-based measures of executive function provide information regarding performance in highly structured environments where the examiner has set the goals and outcomes for the testing session. If performance is low in this optimal, structured testing environment, this might tell us something about potential processing weaknesses in the individual. If performance in this structured environment is at least average and less variable than in unstructured environments, this indicates that a structured environment facilitates performance. Better performance in the standardized assessment context should be taken as an indicator of how well the child would do in the classroom with additional structure and support. The standardized assessment situation has been critiqued for providing a less ecologically valid assessment of how a child performs in 'real' or everyday contexts. However, the standardized assessment situation provides a 'good test' of how well a child will perform under high structure and direction from an examiner. Instead of regarding the 1:1 behavior testing situation in an assessment as a nonecologically valid indicator of behavior, it should be regarded as an indicator of how well performance is ameliorated under highly structured conditions. This is a somewhat novel perspective on the standardized assessment context.

The different information provided by performance-based and rating measures of executive function should allow us to draw prescriptive recommendations for children. The importance of different structured contexts on ADHD behavior has been addressed from the perspective of assessment. For example, the Parental Account of Childhood Symptoms (PACS) interview (Chen et al., 2008; Müller et al., 2011; Taylor, Everitt, et al., 1986, Taylor, Schachar, et al., 1986, 1987) presents parents with questions about behavior during specific situations. Similarly, the Teacher Interview Probe

[2] A somewhat separate but important issue from the perspective of ADHD is that the association between ratings of ADHD severity are significantly correlated with executive function ratings, ranging from $r = .68$ to .91 (Barkley & Murphy, 2010a,b). This raises the question of the degree of overlap between items used to assess ADHD and items used to rate executive function behaviors. This is an important question that is outside the scope of this study.

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

(TIP; Corkum, Andreou, Schachar, Tannock, & Cunningham, 2007) poses parents and teachers with six problem situations. For example, teachers are asked about arrival routines, getting materials ready for lessons, doing group work, doing individual seat work, coming in and settling after morning recess, and getting along with peers. All these situations represent different levels of structure and expectations. Pervasiveness of these symptoms may be apparent across these different contexts, but the structure of the situation may impact the degree to which symptoms of ADHD are expressed. There is good reason to expect that more structure is beneficial for children and youth with ADHD given the effectiveness of behavioral parent training and behavioral classroom management programs (Pelham & Fabiano, 2008). It is thus likely that performance can be ameliorated with intervention strategies increasing degree of structure in the environment for children with ADHD.

Conclusion

Converging evidence supports the conclusion that performance-based and rating measures of executive function assess different aspects of executive function. The administration, task demands, and scoring of these domains of measures are different. Performance-based measures involve considerable structure and direction from the examiner, whereas ratings measures involve very little direction from the examiner. A summary of the empirical work that has examined the association between performance-based and rating measures of executive function demonstrates a very small to modest association between these domains of measures. Theoretical perspectives from the cognitive science literature suggest that performance-based and rating measures of executive function capture different cognitive levels of analysis. Specifically, performance-based measures provide an indication of processing efficiency (the algorithmic mind) and rating measures provide an indication of individual goal pursuit (the reflective mind).

An important implication is that one should not presume that performance-based and ratings measures of executive function capture the same level of analysis, underlying process, or neural substrate. Thus, these measures should not be used interchangeably as parallel measures of executive function in clinical assessments. Both domains of assessment are useful and valuable, but they provide different types of information in the context of clinical assessment.

Supporting information

Additional supporting information is provided along with the online version of this article.

Table S1 Summary of studies examining associations between performance-based measures of executive function and ratings of executive function (Word document)

Please note: Wiley-Blackwell Publishing are not responsible for the content or functionality of any supporting materials supplied by the authors (although this material was peer reviewed by JCPP referees and Editors along with the main article). Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgement

Funding for this study was provided by the Social Sciences and Humanities Research Council (SSHRC) to all three authors.

Correspondence

Maggie Toplak, 126 BSB, Department of Psychology, York University, 4700 Keele St., Toronto, ONT M3J 1P3, Canada; Email: mtoplak@yorku.ca

Key points

- The implications for the practicing clinician are that there is more separability than commonality among performance-based and ratings of executive function. These two classes of measures should not be interpreted as equivalent, interchangeable, or as types or subcategories of one another.
- Performance-based measures of executive function occur under maximal or optimal performance situations and assess the processing efficiency of cognitive abilities under very structured conditions. Rating measures of executive function occur under typical performance situations and assess the extent to which individuals accomplish goal pursuits under unstructured conditions. The former are 'supervisory' and the latter involve 'executive control'.
- Together, these divergent sets of information provide an indication of how well or poorly an individual responds in structured versus unstructured conditions. That is, how well the individual performs when the goals are explicitly laid out versus when the individual must execute his/her own goals without explicit guidance.

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

Areas for future research

- Future research will need to further determine the separable behavioral correlates and outcomes of performance-based and rating measures of executive function.

References

*References marked with an asterisk indicate studies included in the statistical review.

Ackerman, P.L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D.K. Detterman (Ed.), *Current topics in human intelligence* (vol. 4, pp. 1–27). Norwood, NJ: Ablex.

Ackerman, P.L. (1996). A theory of adult development: Process, personality, interests, and knowledge. *Intelligence*, *22*, 227–257.

Ackerman, P.L., & Heggestad, E.D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Alloway, T.P., Gathercole, S.E., & Kirkwood, H.J. (2008). *Working Memory Rating Scale*. London: Pearson Assessment.

American Psychiatric Association. (2000). *Diagnostic and statistical manual for mental disorders* (4th edn rev.). Washington, DC: Author.

Anderson, J.R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, P.J. (2008). Toward a developmental model of executive function. In V. Anderson, R. Jacobs & P.J. Anderson (Eds.), *Executive functions and the frontal lobes* (pp. 3–21). New York: Psychology Press.

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

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

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

Barkley, R.A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology*, *36*, 137–161.

Barkley, R.A., & Murphy, K.R. (2010a). The nature of executive function (EF) deficits in daily life activities in adults with ADHD and their relationship to performance on EF tests. *Journal of Psychopathology and Behavioral Assessment*, *33*, 137–158.

Barkley, R.A., & Murphy, K.R. (2010b). Impairment in occupational functioning and adult ADHD: The predictive utility of executive function (EF) ratings versus EF tests. *Archives of Clinical Neuropsychology*, *25*, 157–173.

Barkley, R.A., & Murphy, K.R. (2011). The nature of executive function (EF) deficits in daily life activities in adults with ADHD and their relationship to performance on EF tests. *ournal of Psychopathology and Behavioral Assessment*, *33*, 137–158.

Barrash, J., Tranel, D., & Anderson, S.W. (2000). Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. *Developmental Neuropsychology*, *2000*, 355–381.

Barry, D., & Petry, N.M. (2008). Predictors of decision-making on the Iowa Gambling Task: Independent effects of lifetime history of substance use disorders and performance on the Trail Making Test. *Brain and Cognition*, *66*, 243–252.

*Bennett, P.C., Ong, B., & Ponsford, J. (2005). Measuring executive dysfunction in an acute rehabilitation setting: Using the dysexecutive questionnaire (DEX). *Journal of the International Neuropsychological Society*, *11*, 376–385.

Biederman, J., Petty, C.R., Fried, R., Black, S., Faneuil, A., Doyle, A.E., … & Faraone, S.V. (2008). Discordance between psychometric testing and questionnaire-based definitions of executive function deficits in individuals with ADHD. *Journal of Attention Disorders*, *12*, 92–102.

*Bodnar, L.E., Prahme, M.C., Cutting, L.E., Denckla, M.B., & Mahone, E.M. (2007). Construct validity of parent ratings of inhibitory control. *Child Neuropsychology*, *13*, 345–362.

Bratman, M.E., Israel, D.J., & Pollack, M.E. (1991). Plans and resource-bounded practical reasoning. In J. Cummins & J. Pollock (Eds.), *Philosophy and AI: Essays at the interface* (pp. 7–22). Cambridge, MA: MIT Press.

Brown, T.E. (2001). *Brown attention deficit disorder scales for children and adolescents* (2nd edn). San Antonio, TX: The Psychological Corporation.

*Brown, T.M., Ris, M.D., Beebe, D., Ammerman, R.T., Oppenheimer, S.G., Yeates, K.O., & Enrile, B.G. (2008). Factors of biological risk and reserve associated with executive behaviors in children and adolescents with spina bifida myelomeningocele. *Child Neuropsychology*, *14*, 118–134.

*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.

Cavallaro, R., Cavedini, P., Mistretta, P., Bassi, T., Angelone, S.M., Ubbiali, A., & Bellodi, L. (2003). Basal-corticofrontal circuits in schizophrenia and obsessive-compulsive disorder: A controlled, double dissociation study. *Biological Psychiatry*, *54*, 437–443.

Chen, W., Zhou, K., Sham, P., Franke, B., Kuntsi, J., Campbell, D., & Asherson, P. (2008). DSM-IV combined type ADHD shows familial association with sibling trait scores: A sampling strategy for QTL linkage. *American Journal of Medical Genetics Part B*, *147B*, 1450–1460.

Clark, L., Manes, F., Antoun, N., Sahakian, B.J., & Robbins, T.W. (2003). The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia*, *41*, 1474–1483.

*Conklin, H.M., Salorio, C.F., & Slomine, B.S. (2008). Working memory performance following paediatric traumatic brain injury. *Brain Injury*, *22*, 847–857.

Corkum, P., Andreou, P., Schachar, R., Tannock, R., & Cunningham, C. (2007). The telephone interview probe: A novel measure of treatment response in children with ADHD. *Educational and Psychological Measurement*, *67*, 169–185.

Cronbach, L.J. (1949). *Essentials of psychological testing*. New York: Harper.

Crone, E.A., Jennings, J.R., & van der Molen, M.W. (2003). Sensitivity to interference and response contingencies in attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *44*, 214–226.

Cummings, J.L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D.A., & Gornbein, J. (1994). The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*, *44*, 2308–2314.

Damasio, A.R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Avon Books.

Damasio, A.R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical*

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

Transactions of the Royal Society of London, Series B, 1413–1420.

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.

Dennett, D.C. (1987). *The intentional stance*. Cambridge, MA: The MIT Press.

Emslie, H.C., Wilson, F.C., Burden, V., Nimmo-Smith, I., & Wilson, B.A. (2003). *Behavioural assessment of the dysexecutive syndrome in children (BADS-C)*. Bury St Edmunds: Thames Valley Test Company.

*Enticott, P.G., Ogloff, J.R.P., & Bradshaw, J.L. (2006). Associations between laboratory measures of executive inhibitory control and self-reported impulsivity. *Personality and Individual Differences*, *41*, 285–294.

Ernst, M., Grant, S.J., London, E.D., Contoreggi, C.S., Kimes, A.S., & Spurgeon, B.A. (2003). Decision making in adolescents with behavior disorders and adults with substance abuse. *American Journal of Psychiatry*, *160*, 33–40.

Gioia, G.A., Espy, K.A., & Isquith, P.K. (2003). *Behavior rating inventory of executive function – Preschool version*. Odessa, Florida: Psychological Assessment Resources.

Gioia, G.A., Isquith, P.K., Guy, S.C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function*. Odessa, Florida: Psychological Assessment Resources.

Gioia, G.A., Isquith, P.K., & Kenealy, L.E. (2008). Assessment of behavioral aspects of executive function. In V. Anderson, R. Jacobs & P.J. Anderson (Eds.), *Executive functions and the frontal lobes* (pp. 179–202). New York: Psychology Press.

Grace, J., & Malloy, P.F. (2001). *Frontal Systems Behavior Scale (FrSBe): Professional Manual*. Lutz, FL: Psychological Assessment Resources.

Guy, S.C., Isquith, P.K., & Gioia, G.A. (2005). *Behavior rating inventory of executive function –Self report version*. Odessa, Florida: Psychological Assessment Resources, Inc.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., & Curtis, G. (1993). *Wisconsin Card Sorting Test (WCST) manual, revised and expanded*. Odessa, FL: Psychological Assessment Resources.

*Hummer, T.A., Kronenberger, W.G., Wang, Y., Dunn, D.W., Mosier, K.M., Kalmin, A.J., & Mathews, V.P. (2010). Executive functioning characteristics associated with ADHD comorbidity in adolescents with Disruptive Behavior Disorders. *Journal of Abnormal Child Psychology*, *39*, 11–19.

Jensen, A.R., & Rohwer, W.D. Jr. (1966). The Stroop color-word test: A review. *Acta Psychologica*, *25*, 36–93.

Kertesz, A., Davidson, W., & Fox, H. (1997). Frontal behavioral inventory: Diagnostic criteria for frontal lobe dementia. *Canadian Journal of Neurological Sciences*, *24*, 29–36.

Kester, H.M., Sevy, S., Yechiam, E., Burdick, K.E., Cervellione, K.L., & Kumra, S. (2006). Decision-making impairments in adolescents with early-onset schizophrenia. *Schizophrenia Research*, *85*, 113–123.

Labudda, K., Frigge, K., Horstmann, S., Aengenendt, J., Woermann, F.G., Ebner, A., … & Brand, M. (2009). Decision making in patients with temporal lobe epilepsy. *Neuropsychologia*, *47*, 50–58.

Lamm, C., Zelazo, P.D., & Lewis, M.D. (2006). Neural correlates of cognitive control in childhood and adolescence: Disentangling the contributions of age and executive function. *Neuropsychologia*, *44*, 2139–2148.

Lawrence, N.S., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A., & Phillips, M.L. (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology*, *20*, 409–419.

MacLeod, C.M. (1991). Half a century of research on the Stroop Effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.

Mahmut, M.K., Homewood, J., & Stevenson, R.J. (2008). The characteristics of non-criminals with high psychopathy traits: Are they similar to criminal psychopaths? *Journal of Research in Personality*, *42*, 679–692.

*Mahone, E.M., Cirino, P.T., Cutting, L.E., Cerrone, P.M., Hagelthorn, K.M., Hiemenz, J.R., … & Denckla, M.B. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*, 643–662.

Malloy, P., & Grace, J. (2005). A review of rating scales for measuring behavior change due to frontal systems damage. *Cognitive and Behavioral Neurology*, *18*, 18–27.

*Mangeot, S., Armstrong, K., Colvin, A.N., Yeates, K.O., & Taylor, H.G. (2002). Long-term executive function deficits in children with traumatic brain injuries: Assessment using the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychology*, *8*, 271–284.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Martel, M.M., & Nigg, J.T. (2006). Child ADHD and personality/temperament traits of reactive and effortful control, resiliency, and emotionality. *Journal of Child Psychology and Psychiatry*, *47*, 1175–1183.

Matthews, G., Zeidner, M., & Roberts, R.D. (2002). *Emotional intelligence: Science & myth*. Cambridge, MA: MIT Press.

*McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*, 495–505.

Miley, W.M., & Spinella, M. (2006). Correlations among measures of executive function and positive psychological attributes in college students. *The Journal of General Psychology*, *133*, 175–182.

Miyake, A., & Friedman, N.P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*, 8–14.

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

Müller, U.C., Asherson, P., Banaschewski, T., Buitelaar, J.K., Ebstein, R.P., Eisenberg, J., & Steinhausen, H.C. (2011). The impact of study design and diagnostic approach in a large multicentre ADHD study. Part 1: ADHD symptom patterns. *BMC Psychiatry*, *11*, 54.

Nakamura, M., Nestor, P.G., Levitt, J.J., Cohen, A.S., Kawashima, T., Shenton, M.E., & McCarley, R.W. (2008). Orbitofrontal volume deficit in schizophrenia and thought disorder. *Brain*, *131*, 180–195.

Newell, A. (1982). The knowledge level. *Artificial Intelligence*, *18*, 87–127.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

*Niendam, T.A., Horwitz, J., Bearden, C.E., & Cannon, T.D. (2007). Ecological assessment of executive dysfunction in the psychosis prodrome: A pilot study. *Schizophrenia Research*, *93*, 350–354.

Nigg, J.T. (2006). *What causes ADHD? Understanding what goes wrong and why*. New York: Guilford Press.

Nigg, J.T., Stavro, G., Ettenhofer, M., Hambrick, D.Z., Miller, T., & Henderson, J.M. (2005). Executive functions and ADHD in adults: Evidence for selective effects on ADHD symptom domains. *Journal of Abnormal Psychology*, *114*, 706–717.

*Norris, G., & Tate, R.L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation*, *10*, 33–45.

*Odhuba, R.A., van den Broek, M.D., & Johns, L.C. (2005). Ecological validity of measures of executive functioning. *British Journal of Clinical Psychology*, *44*, 269–278.

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.

*Parrish, J., Geary, E., Jones, J., Seth, R., Hermann, B., & Seidenberg, M. (2007). Executive functioning in childhood epilepsy: Parent-report and cognitive assessment. *Developmental Medicine and Child Neurology*, *49*, 412–416.

Pelham, W.E., & Fabiano, G.A. (2008). Evidence-based psychosocial treatments for Attention-Deficit/Hyperactivity Disorder. *Journal of Clinical Child and Adolescent Psychology*, *37*, 184–214.

Pennington, B.F. (2002). *The development of psychopathology: Nature and nurture*. New York: Guilford.

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

Pollock, J.L. (1995). *Cognitive carpentry: A blueprint for how to build a person*. Cambridge, MA: MIT Press.

*Riccio, C.A., Hall, J., Morgan, A., Hynd, G.W., & Gonzalez, J.J. (1994). Executive function and the Wisconsin Card Sorting Test: Relationship with behavioral ratings and cognitive ability. *Developmental Neuropsychology*, *10*, 215–229.

Roth, R.M., Isquith, P.K., & Gioia, G.A. (2005). *Behavior rating inventory of executive function- Adult version*. Lutz, Florida: Psychological Assessment Resources, Inc.

Sagvolden, T., Johansen, E.B., Aase, H., & Russell, V.A. (2005). A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behavioral and Brain Sciences*, *28*, 397.

Salthouse, T.A., Atkinson, T.M., & Berish, D.E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566–594.

Scheres, A., Oosterlaan, J., Geurts, H., Morein-Zamir, S., Meiran, N., Schut, H., … & Sergeant, J.A. (2004). Executive functioning in boys with ADHD: Primarily an inhibition deficit? *Archives of Clinical Neuropsychology*, *19*, 569–594.

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

Shallice, T. (1990). *From neuropsychology to mental structure*. New York: Cambridge University Press.

*Shuster, J., & Toplak, M.E. (2009). Executive and motivational inhibition: Associations with self-report measures related to inhibition. *Consciousness and Cognition*, *18*, 471–480.

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

Sonuga-Barke, E.J.S., Dalen, L., & Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 1335–1342.

Spinella, M. (2005). Self-rated executive function: Development of the executive function index. *International Journal of Neuroscience*, *115*, 649–667.

Stanovich, K.E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stanovich, K.E. (2009a). Distinguishing the reflective, algorithmic, and autonomous minds: It is time for a tri-process theory? In J.St.B.T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55–88). New York: Oxford University Press.

Stanovich, K.E. (2009b). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Stanovich, K.E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.

Stanovich, K.E., West, R.F., & Toplak, M.E. (2011). Intelligence and rationality. In R.J. Sternberg & S.B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 784–826), New York: Cambridge University Press.

Strauss, E., Sherman, E.M.X., & Spreen, O. (2006). *A compendium of neuropsychological tests* (3rd edn). New York: Oxford University Press.

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

Taylor, E., Everitt, B., Thorley, G., Schachar, R., Rutter, M., & Wieselberg, M. (1986). Conduct disorder and hyperactivity: II. A cluster analytic approach to the identification of a behavioural syndrome. *British Journal of Psychiatry*, *149*, 768–777.

Taylor, E., Schachar, R., Thorley, G., Wieselberg, H.M., Everitt, B., & Rutter, M. (1986). Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *British Journal of Psychiatry*, *149*, 760–767.

Taylor, E., Schachar, R., Thorley, G., Wieselberg, H.M., Everitt, B., & Rutter, M. (1987). Which boys respond to stimulant medication? A controlled trial of methylphenidate in boys with disruptive behaviour. *Psychological Medicine*, *17*, 121–143.

Thorell, L.B., Eninger, L., Brocki, K.C., & Bohlin, G. (2010). Childhood Executive Function Inventory (CHEXI): A promising measure for identifying young children with ADHD? *Journal of Clinical and Experimental Neuropsychology*, *32*, 38–43.

Thorell, L.B., & Nyberg, L. (2008). The Childhood Executive Functioning Inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology*, *33*, 536–552.

*Toplak, M.E., Bucciarelli, S., Jain, U., & Tannock, R. (2008). Executive functions: Performance-based measures and the Behavior Rating Inventory of Executive Function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*, 53–72.

Toplak, M.E., Jain, U., & Tannock, R. (2005). Executive and motivational processes in adolescents with attention-deficit-hyperactivity disorder (ADHD). *Behavioral and Brain Functions*, *1*, 8. Available from: http://www.behavioralandbrainfunctions.com/content/1/1/8.

Toplak, M.E., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K.E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making*, *20*, 103–124.

*Vriezen, E.R., & Pigott, S.E. (2002). The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology*, *8*, 296–303.

Williams, B.R., Ponesse, J.S., Schachar, R.J., Logan, G.D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology*, *35*, 205–213.

Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H., & Evans, J.J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome (BADS)*. UK: Thames Valley Test Company.

*Wilson, B.A., Evans, J.J., Emslie, H., Alderman, N., & Burgess, P. (1998). The development of an ecologically valid test for assessing patients with a dysexecutive syndrome. *Neuropsychological Rehabilitation*, *8*, 213–228.

Accepted for publication: 28 August 2012
Published online: 11 October 2012

© 2012 The Authors. Journal of Child Psychology and Psychiatry © 2012 Association for Child and Adolescent Mental Health.