sciendo

SJCAPP

Scandinavian Journal of Child and Adolescent Psychiatry and Psychology
Vol. 12:1-9 (2024) DOI 10.2478/sjcapp-2024-0001

Research Article **Open Access**

Executive functioning in children with ADHD Investigating the cross-method correlations between performance tests and rating scales

Kristoffer Dalsgaard Olsen[1]*, Denis Sukhodolsky[2], Aida Bikic[3]

[1] Child and Youth Psychiatric outpatient clinic, Region of Southern Denmark, Aabenraa, Denmark
[2]Child Study Centre, Yale University School of Medicine, New Haven, United States
[3]Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

***Corresponding author:** kristofferdalsgaard@gmail.com

Abstract

Objective: Replicated evidence shows a weak or non-significant correlation between different methods of evaluating executive functions (EF). The current study investigates the association between rating scales and cognitive tests of EF in a sample of children with ADHD and executive dysfunction.
Method: The sample included 139 children (aged 6-13) diagnosed with ADHD and executive dysfunctions. The children completed subtests of the Cambridge Neuropsychological Test Automated Battery (CANTAB). Parents completed the Behavior Rating Inventory of Executive Function (BRIEF) and the Children's Organizational Skills Scale (COSS).
Analysis: Pairwise Spearman correlations were calculated between the composite and separate subscales of cognitive tests and rating scales. In secondary analyses, pairwise Spearman correlations were conducted between all composite scales and subscales, stratified by child sex and child ADHD subtype.
Results: The correlation analyses between composite scores yielded no significant correlations. The results when comparing CANTAB TO and BRIEF GE are r=-.095, p=.289, and r=.042, p=.643 when comparing CANTAB TO and COSS TO. The analyses between all composite scales and subscales found one significant negative correlation (r=-.25, p<.01). There are significant cross-method differences when stratified by the ADHD-Inattentive subtype, showing significant negative correlations (moderate) between CANTAB and BRIEF composite (r=-.355, p=.014) and subscales.
Discussion: It is possible that the different methods measure different underlying constructs of EF. It may be relevant to consider the effects of responder bias and differences in ecological validity in both measurement methods.
Conclusion: The results found no significant correlations. The expectation in research and clinical settings should not be to find the same results when comparing data from cognitive tests and rating scales. Future research might explore novel approaches to EF testing with a higher level of ecological validity, and designing EF rating scales that capture EF behaviors more so than EF cognition.

Keywords: ADHD, executive functions, cognition, behavior, cross-methodological, ecological validity, cognitive test, performance test, rating scale, BRIEF, CANTAB, COSS

Introduction

Executive functions (EF) represent higher-order neurocognitive processes involved in the regulation and direction of thoughts and behaviors, relevant to problem-solving, goal setting, and following plans (1). From the multitude of specific EF subtypes, three core EFs are identified: *1) Inhibition*: the ability to exert behavioral self-control and cognitive interference control when confronted with external/internal stimuli. *2) Working memory*: the ability to retain relevant information in one's mind while using it to solve a present problem. *3) Cognitive flexibility*: the ability to change perspective and/or approach when confronted with changes such as novel information, rules, or changes in relevance (1).

EF is of central importance to a person's daily functioning and is an important factor in the organization of actions to reach present and future goals. EF plays a central role in the current understanding of ADHD, and the basic dysfunctions underlying the disorder, as well as in other psychiatric disorders (2-5).

Even though EF play a central role in ADHD, and individuals with ADHD score significantly lower than control groups on measures of EF (6, 7), research also shows that the relationship between ADHD and EF is not well understood. Evidence shows that not all children with ADHD express deficits in areas of EF, although children with ADHD show a higher risk of these deficits (8-10). High levels of cognitive heterogeneity are present in children with ADHD (11, 12). Evidence indicates only weak correlations between different methods of evaluating EF. Toplak et al. (13, 14), Krieger et Amador-Campos (6), Soto et al. (15), and others have found some modest, primarily non-significant or weak correlations between cognitive tests and rating scales measuring EF. In a meta-analysis conducted by Toplak et al. (14) examining the overall cross-method correlation in 20 individual studies, the overall median correlation between EF test and rating scales was only r=.19.

Toplak et al. (13) find that the rating scale Behavior Rating Inventory of Executive Function (BRIEF) expresses a superior level of predicting ADHD core symptoms. They argue that the cause of the non-significant levels of correlations between cognitive tests and rating scales may be related to the gap in ecological validity, which refers to differences in the assessment methods ability to predict behavior in real-world settings. Toplak et al. (13) argue that the clinical setting of cognitive tests produces poor ecologically valid information compared to information gathered from rating scales which are evaluated by parents and teachers, who experience the behavior of the child in their daily life. Toplak et al. (14) argue that cognitive tests cannot be used to predict performance in goal-directed behaviors in real-life settings, and go on to hypothesize that rating scales, on the other hand, might reflect mental constructs pertaining to the successful pursuit of goals.

Ledochowski, Andrade & Toplak (16) further argue that the implementation of a test design that seeks to emulate naturalistic EF tasks and settings might produce more ecologically valid EF test results. Ledochowski, Andrade & Toplak (16) propose the use of Unstructured Performance-Based Task (UPT), a test design in which the test material and examiner impose minimal structure on the test subject.

Soto et al. (15), on the other hand, argued that the research design of earlier studies, such as Toplak et al. (13, 14), has an impact on the interpretation of these studies. Parent ratings of ADHD symptoms were used to evaluate the predictive qualities of rating scales and cognitive tests, and the outcome was determined by mono-informant and mono-method biases. Soto et al.'s (15) research indicates that:

1. EF tests outperform EF ratings in predicting academic performance (regarding academic test results defined by The Kaufman Test of Educational Achievement), while EF tests and EF ratings show similar levels of correlation with academic rating results (defined by The Academic Performance Rating Scale).
2. A strong correlation exists between academic performance tests and academic ratings.
3. EF test and EF rating results showed a significant but only modest correlation.

Based on these complex results, Soto et al. (15) conclude that EF tests and rating scales must reflect different underlying constructs or expressions of EF (cognitive and behavioral aspects). Soto et al. (15) argue that the use of rating scales, which adopt alternative terms and constructs that refer to specific behavioral skills, such as organizational, time management, and planning skills, may help to distinguish the different but related areas of executive function (i.e., cognitive, and behavioral expressions of EF skills).

As presented in previous studies of EF and ADHD, the results showed weak or no correlations between cognitive tasks and ratings of EF in children with ADHD. However, the studies included children with various degrees of EF dysfunctions, as shown children with ADHD show great heterogeneity of cognitive functions. The present study aims to investigate the correlations between EF cognitive tests and parent rating scales in a group of children with ADHD, who also specifically show significant impairments of EF.

Therefore, the current study aims to investigate a clinically meaningful subgroup of children with both ADHD and significant executive dysfunctions defined by high ratings on the Behavior Rating Inventory of Executive Function (BRIEF) organize subscale. Organizational skills are a relatively new area of neurocognitive and clinical research but shows promise as an area of intervention research. The current study employs cognitive tests and rating scales that measure a broad array of EF, and questionnaires designed to measure organizational and planning skills specifically. The study may clarify how best to design and interpret the results when using different modalities to measure EF.

Aim

To investigate cross-method correlation between cognitive tests and parent rating scales of EF in children having both ADHD and significant executive dysfunctions in planning and organization skills.

Ethical considerations

This study is based on baseline data from a randomized controlled trial (RCT) investigating the effect of "Organizational Skills Training for Children with ADHD" (17). The number of participants was determined with a sample size calculation based on the primary outcome of the RCT, for more details please see the protocol (17). The data was collected between June 2017 and November 2022. A child could only participate in the trial if the written consent of all legal guardians was obtained. Danish law requires written consent by the child above the age of 15 years. All participating children gave verbal assent for participation. The trial has been approved by the Danish Data Protection Agency (ID 17/7467) and the Regional Scientific Ethical Committees for Southern Denmark (S-20160180).

Sample characteristics

The current study investigates a relatively homogeneous sample of children diagnosed with ADHD. All participants were recruited from outpatient clinics in the Southern Region of Denmark in an intervention trial of Organizational Skills training (17) using the following inclusion criteria: All ADHD diagnoses were verified using the Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADS) diagnostic screening interview (18). All children show an IQ >80 using the Reynolds Intellectual Screening Test (RIAS) (19) and BRIEF parent rating scale scores on or above the clinical cut-off of 60 on the plan/organize subscale. Children with an autism spectrum disorder were excluded from the study. Only data collected at baseline were used in this paper.

Measures

Behavior Rating Inventory of Executive Function - BRIEF is a rating scale using a three-point rating system (never/sometimes/often), designed to measure the executive function of children (age 5 to 18). It contains 86 items, eight clinical scales, and is comprised of a parent and a teacher questionnaire. BRIEF's internal consistency (αs = .80–.98) and test/retest reliability (rs = .82 for parents, .88 for teachers) is good, the inter-tester reliability between parent and teacher rating (r = .32) is, however, low (20-22).

Children's Organizational Skills Scale - COSS (23) is a rating scale designed to measure specific domains in children's behavioral abilities at home and school life concerning skills in Organization, Time Management, and Planning (OTMP skills). The COSS has 58 items (parents) and 35 (teachers) using a 4-point scale (hardly ever or never/sometimes/much of the time/just about all the time). COSS shows excellent internal consistency alphas for Parent (αs = .98) and Teacher (αs = .97) items alike, as well as test/retest reliability (r = .99 for parents and .94 for teachers). COSS shows good interrater consistency (mean parent to teacher being r = .69, p < .001). The COSS scale furthermore measures specific organizational behavior, in line with Soto et al.'s (15) argument for the need of rating scales that target EF behavior specifically.

Cambridge Neuropsychological Test Automated Battery – CANTAB is a computer-administered test battery containing high levels of precise, sensitive, and objective measurements concerning cognitive functions, and is often used in the research of EF (6, 7, 13). In the present study, the following tests were used: One Touch Stockings of Cambridge (OTS) an executive functions measure addressing planning. The Spatial Working Memory Test (SWM) a nonverbal working memory measure and, the Stop Signal Task (SST) measuring impulsivity.

Subscale and composite level of EF

To take the complex and interconnected structure of EF into account, the present study applies a correlation analysis that encompasses both subscale levels and composite levels of EF – both in respect to the rating scales and cognitive test data. Subscales and composite variables from cognitive test (CANTAB): from the Spatial Working Memory (SWM) test, which measures working memory, the variable SWM Between Search Errors (SW) is applied. From the One Touch Stockings of Cambridge (OTS) test, which measures planning, the variable OTS Problems Solved on First Choice (OT) is applied. From the Stop Signal Task (SST) test, which measures inhibition, the variables SST Stop Signal Reaction Time, SST Directions Errors: Go Trials, and SST Directions Errors: Stop Trials are all weighted and applied as a single measure: SST weighted variable (SS). This was done by adding z-standardized scores of the OTS CANTAB items with weights 2:1:1. A Global cognitive composite score has been created by combining the present subscales into one combined score: CANTAB Composite Total Score (TO). The composite CANTAB TO scale was computed as a sum score of z-standardized OT, SW and SS scores.

From the BRIEF rating scale, the following subscales and composite variables are applied: Global Executive Composite (GE), Emotional Control (EM), Shift (SH), Initiate (IN), Organization of Materials (MA), Inhibition (IH), Working Memory (ME), Monitor (MO), Plan/Organize (PL).

From the COSS rating scale, the following variables are applied: Memory and Materials management (MM), Task Planning (TP), Organized Actions (OA) and Total score (TO).

Statistical Analyses

As primary analyses, pairwise Spearman correlations were calculated between the three composite scales, BRIEF GE, COSS TO, and CANTAB TO, and for the same-variable subscales. In secondary analyses, pairwise Spearman correlations were conducted between all composite and subscales of BRIEF, COSS, and CANTAB. In the interpretation of the secondary analysis, a p-value of .01 indicates a significant finding, to take the multiple testing into account. Further secondary analyses considered pairwise Spearman correlations stratified on child sex, and child ADHD subtype (excluding children with hyperactive subtype due to an insufficient number of cases). The t-scores of COSS scales are only defined for children aged 8 years or above, however, the study sample included 10 children of age 6-7 years. In a sensitivity analysis, children of ages 6-7 years were allocated COSS t-scores corresponding to age 8 years, and pairwise Spearman correlations between all composite and subscales of BRIEF, COSS, and CANTAB were computed, including COSS t-scores of children aged 6-7 years. The analysis found no significant differences between children ages 6-7 and children ages 8-13 on the COSS scale results.

TABLE 1. Participant characteristics (N=139).

Sex, n (%)	
Girl	31 (22%)
Boy	108 (78%)
Child age, mean *(SD)* [range]	10.2 *(1.7)* [6-13]
ADHD subtype, n (%)	
Predominately inattentive	52 (37%)
Hyperactive	5 (4%)
Combined	82 (59%)
Comorbidity, n (%)	
None	97 (70%)
≥ 1	42 (30%)
IQ, mean *(SD)* [range]	97.5 *(7.8)* [81-23]
IQ tertiles, n (%)	
Lower (81-93)	46 (33%)
Middle (94-100)	41 (30%)
Higher (101-123)	52 (37%)
BRIEF GE, mean *(SD)* [range]	72.6 *(7.2)* [53-88]
COSS TO, mean *(SD)* [range]	67.6 *(6.4)* [54-80]

BRIEF GE: Behavior Rating Inventory of Executive Function Global Executive Composite; COSS TO: Childrens Organizational Skills Scale Total scale; CANTAB TO: Cambridge Neuropsychological Test Automated Battery Composite total score

All analyses were performed in STATA 17.0 (Statacorp, Texas, USA).

Results

Descriptive statistics were computed on the participant characteristics: child sex, child age, ADHD subtype. See Table 1 for more details.

Primary outcomes

The pairwise correlation analyses between composite scores yielded no statistically significant correlations when comparing different methods of assessing EF, as shown in table 2. The result when comparing CANTAB TO and BRIEF GE is r=-.095, p=.289), and r=.042, p=.643 when comparing CANTAB TO and COSS TO. The results do however show a significant and strong correlation when comparing BRIEF and COSS composite scores (r=.526, p<.001).

TABLE 2. Pairwise correlations between composite scores: r (*p*-value)

	BRIEF GE *n = 129*	*COSS TO* *n = 124*	*CANTAB TO* *n = 137*
BRIEF GE	1		
COSS TO	.526 (<0.001)	1	
CANTAB TO	-.095 (0.289)	.042 (.643)	1

BRIEF GE: Behavior Rating Inventory of Executive Function Global Executive Composite; COSS TO Childrens Organizational Skills Scale Total scale; CANTAB TO; Cambridge Neuropsychological Test Automated Battery Composite total score.

The pairwise correlation analysis between the same subscale-items shows no statistically significant correlations when comparing different methods of assessing EF, as shown in Table 3. It is important to note that CANTAB-OT is the only item that scores positive; a high score represent a positive result and a low score a negative. All other items score negative. Therefor all results related to CANTAB-OT must be interpreted as positive-negative inverted, such that a negative correlation is a positive correlation and vice versa. For the planning subscale: CANTAB-OT and BRIEF-PL, the result is (r = -.10), and for the CANTAB-OT and COSS-TP, the result is (r = -.08). For the working memory subscale: The CANTAB-SW and BRIEF-ME result is (r = .09), and the CANTAB-SW and COSS-MM is (r = -.05). For the Inhibition subscale: The CANTAB-SS and BRIEF-IH result is (r = -.01), and the CANTAB-SW and COSS-OA is (r = .05). Comparing the results from the pairwise analysis of items from COSS and BRIEF, the results range from (r = .24, p = .01) BRIEF-PL and COSS-TP, (r = .34, p = .001) BRIEF-ME and COSS-MM, (r = .03) BRIEF-IH and COSS-OA.

TABLE 3. Means and SDs of subscales, and pairwise correlations between subscales.

	BRIEF, n = 129									COSS, n = 124				CANTAB, n = 137			
	GE	*EM*	*SH*	*IN*	*MA*	*IH*	*ME*	*MO*	*PL*	*TO*	*TP*	*OA*	*MM*	*TO*	*OT*	*SW*	*SS*
Mean	72.6	64.8	71.3	65.6	63.5	67.3	73.7	66.8	70.5	67.6	66.8	63.3	65.5	-.003	.000	.000	-.000
(SD)	*(7.2)*	*(11.9)*	*(11.6)*	*(8.1)*	*(7.3)*	*(10.2)*	*(6.3)*	*(7.5)*	*(6.9)*	*(6.4)*	*(9.2)*	*(5.1)*	*(9.2)*	*(.660)*	*(1.004)*	*(1.000)*	*(.894)*
BRIEF																	
GE	1																
EM	.77***	1															
SH	.75***	.68***	1														
IN	.61***	.30***	.37***	1													
MA	.44***	.17^{NS}	.12^{NS}	.22*	1												
IH	.73***	.60***	.47***	.24**	.21*	1											
ME	.60***	.30***	.32***	.45***	.29**	.29***	1										
MO	.66***	.40***	.37***	.35***	.26**	.54***	.23**	1									
PL	.66***	.25**	.35***	.56***	.39***	.22*	.54***	.44***	1								
COSS																	
TO	.53***	.33***	.32***	.26**	.37***	.31***	.42***	.30**	.46***	1							
TP	.27**	.20*	.23*	.17^{NS}	-.07^{NS}	.15^{NS}	.32***	.11^{NS}	.24**	.63***	1						
OA	.23*	.10^{NS}	.04^{NS}	.06^{NS}	.35***	.03^{NS}	.18^{NS}	.26**	.30**	.61***	.21*	1					
MM	.53***	.38***	.31***	.25*	.22*	.41***	.34***	.34***	.41***	.75***	.40***	.29**	1				
CANTAB																	
TO	-.09^{NS}	-.15^{NS}	-.11^{NS}	-.07^{NS}	.09^{NS}	.11^{NS}	-.03^{NS}	-.19*	.14^{NS}	.04^{NS}	.01^{NS}	.18*	-.06^{NS}	1			
OT	.11^{NS}	.20*	.06^{NS}	-.05^{NS}	-.03^{NS}	.25**	-.03^{NS}	.14^{NS}	-.10^{NS}	.02^{NS}	-.08^{NS}	-.22*	.09^{NS}	.74***	1		
SW	.04^{NS}	-.00^{NS}	-.01^{NS}	-.13^{NS}	.18*	.05^{NS}	.09^{NS}	-.06^{NS}	.09^{NS}	.07^{NS}	-.03^{NS}	.10^{NS}	-.05^{NS}	.69***	-.27**	1	
SS	-.12^{NS}	-.10^{NS}	-.17^{NS}	-.05^{NS}	-.04^{NS}	-.01^{NS}	-.06^{NS}	-.18*	.09^{NS}	.04^{NS}	.04^{NS}	.05^{NS}	.05^{NS}	.62***	-.22*	.11^{NS}	1

**p < 0.05, **p < 0.01, ***p < 0.001, NS: not significant (p ≥ 0.05)*

BRIEF: Behavior Rating Inventory of Executive Function; GE: Global Executive Composite; EM: Emotional Control; SH: Shift; IN: Initiate; MA: Organization of Materials; IH: Inhibition; ME: Working Memory; MO: Monitor; PL: Plan/Organize. COSS: Childrens Organizational Skills Scale; TO : Total scale, TP: Task Planning, OA: Organized Actions, MM: Memory and Materials Management; CANTAB: Cambridge Neuropsychological Test Automated Battery; TO: Composite total score, OT: Problems Solved on First Choice; SW: Between Search Errors, SS: weighted variables (inhibition).

Secondary Outcomes

The pairwise correlation between all composites and subscales, when comparing test and rating scale results, finds seven weak correlations (i.e., < r = .3). From the seven possible correlations, four are negatively correlated. These correlations concern the CANTAB test items Planning and Inhibition and the BRIEF rating scale items Emotional control, Inhibition, and Monitoring. Three of the possible correlations are positive. These correlations concern the CANTAB test item Working memory and the BRIEF rating scale item Monitoring, and both the CANTAB test item Planning and the CANTAB Composite total score and the COSS rating scale item Organized actions. All but one correlation falls under the chosen threshold of p = .01, the one correlation with a p-value of .01 is CANTAB OT and BRIEF IH

Subtype – sex

The results indicate no statistically significant cross-methodological differences when stratified for sex, see table 4. The lack of significant correlation between BRIEF and COSS in girls; r = .217 (p = .268) compared to boys; r = .613 (p = <.001) is probably related to the difference in sample size: Girls, n = 30 Boys, n = 105.

TABLE 4. Pairwise correlations between composite scores, stratified on sex: r (*p*-value).

	BRIEF GE n = 129	*COSS TO n = 124*	*CANTAB TO n = 137*
Girls (n = 30)			
BRIEF GE	1		
COSS TO	.217 (.268)	1	
CANTAB TO	-.007 (.971)	.285 (.133)	1
Boys (n= 105)			
BRIEF GE	1		
COSS TO	.613 (<.001)	1	
CANTAB TO	-.052 (.618)	-.004 (.969)	1

BRIEF GE: Behavior Rating Inventory of Executive Function Global Executive Composite; COSS TO: Childrens Organizational Skills Total scale; CANTAB TO: Cambridge Neuropsychological Test Automated Battery Composite total score.

Subtypes – ADHD (I and C)

The results indicate statistical cross-methodological differences when stratified for ADHD sub-type, see table 5. The participants characterized by the ADHD-I (predominantly inattentive) subtype express a significant negative correlation (moderate, r = .3-.5) when comparing CANTAB total scale and BRIEF Global executive composite (r=-.355, p=.014). Table 5 Pairwise correlations between composite scores, stratified on ADHD subtype: r (p-value)

The same result continues in the pairwise correlations analysis (see table 6) between all variables (composite and subscales) when solely analyzing results from participants defined with the primarily inattentive subtype. The results identify 10 significant correlations, two are weak, eight are of moderate strength and nine of the 10 possible correlations are negative correlations. These correlations primarily concern the CANTAB item Planning and the BRIEF items Emotional Control, Inhibition and the CANTAB item concerning working memory and the BRIEF items Initiate and Monitor.

TABLE 5. Pairwise correlations between composite scores, stratified on ADHD subtype: r (p-value).

	BRIEF GE n = 129	*COSS TO n = 124*	*CANTAB TO n = 137*
Predominantly inattentive (n = 52)			
BRIEF GE	1		
COSS TO	.653(<.001)	1	
CANTAB TO	-.355 (.014)	-.105 (.482)	1
Combined (n = 82)			
BRIEF GE	1		
COSS TO	.435 (<.001)	1	
CANTAB TO	-.030 (.796)	.086 (.474)	1

BRIEF GE: Behavior Rating Inventory of Executive Function Global Executive Composite; COSS TO: Childrens Organizational Skills Total scale; CANTAB TO: Cambridge Neuropsychological Test Automated Battery Composite total score.

Discussion

Primary outcomes

The results from the pairwise correlation analysis, when comparing the test and rating scale measuring EF, are in line with earlier findings (6, 13-15). There are no significant correlations between the same variables of rating scales and cognitive tests, and only weak correlations between all variables are found.

Our most promising finding is the significant but weak correlation between the CANTAB-OT measure, which concerns Planning, and the COSS variable concerning Organized Action (r = -.22, p = .05). These items concern similar underlying constructs. In isolation, it is possible to interpret the finding as indicating a significant correspondence between the cognitive ability to construct mental plans and the observable ability to organize one's actions. However, the correlation is only weak, and as evident in the following discussion, the majority of non-significant and negative correlations overshadow this finding in general. The significant correlations could very well represent random findings due to multiple comparisons.

TABLE 6. Pairwise correlations between CANTAB subscales and BRIEF and COSS subscales for children with predominantly inattentive ADHD subtype.

	BRIEF, n = 47									COSS, n = 47			
	GE	*EM*	*SH*	*IN*	*MA*	*IH*	*ME*	*MO*	*PL*	*TO*	*TP*	*OA*	*MM*
CANTAB (n = 52)													
TO	-.35*	-.43**	-.33*	-.26^{NS}	-.03^{NS}	-.24^{NS}	-.10^{NS}	-.36*	.04^{NS}	-.10^{NS}	0.03^{NS}	-0.04^{NS}	-0.23^{NS}
OT	.34*	.41**	.22^{NS}	-.00^{NS}	.09^{NS}	.39**	.08^{NS}	.29^{NS}	.09^{NS}	.12^{NS}	-0.18^{NS}	-0.04*	0.25^{NS}
SW	-.25^{NS}	-.27^{NS}	.18^{NS}	-.39**	.08^{NS}	-.11^{NS}	-.13^{NS}	-.30*	.01^{NS}	-.16^{NS}	-0.08^{NS}	-0.13^{NS}	-0.23^{NS}
SS	-.12^{NS}	-.18^{NS}	-.28^{NS}	-.13^{NS}	-.04^{NS}	.02^{NS}	.01^{NS}	-.14^{NS}	.18^{NS}	.11^{NS}	-0.06^{NS}	0.02^{NS}	.08^{NS}

**p < 0.05, **p < 0.01, ***p < 0.001, NS: not significant (p ≥ 0.05)*

BRIEF: Behavior Rating Inventory of Executive Function; GE: Global Executive Composite; EM: Emotional Control; SH: Shift; IN: Initiate; MA: Organization of Materials; IH: Inhibition; ME: Working Memory; MO: Monitor; PL: Plan/Organize. COSS: Childrens Organizational Skills Scale; TO : Total scale, TP: Task Planning, OA: Organized Actions, MM: Memory and Materials Management; CANTAB: Cambridge Neuropsychological Test Automated Battery; TO: Composite total score, OT: Problems Solved on First Choice; SW: Between Search Errors, SS: weighted variables (inhibition).

Secondary outcomes

The analysis produced a host of non-significant correlations and one significant negative correlation. The negative correlation concerns the CANTAB-OT planning item and the BRIEF-Inhibition. Similar results, i.e., many non-significant and negatively correlated correlations, are presented by Toplak et al. (13, 14) and Krieger & Amador-Campos (6).

Similarities and differences in ADHD sub-types

The results show statistical similarities and differences between the combined group of participants (defined as ADHD-H and ADHD-C) and the ADHD-I group. Both groups show negative correlations between the CANTAB item measuring Planning and the BRIEF items measuring Emotional Control and Inhibition. However, both groups also show positive correlations with the CANTAB item measuring Planning and the item Organized Actions on COSS.

Specific to the combined group is the positive correlation between the CANTAB item measuring Working Memory and the item Organization of Materials on BRIEF, and a negative correlation between the CANTAB item measuring Inhibition and the item Monitoring on BRIEF.

The primarily inattentive group (ADHD-I), on the other hand, showed a specific negative correlation between the CANTAB item measuring Working Memory and the BRIEF items measuring Initiate and Monitoring. Earlier studies have found interesting but inconclusive differences between ADHD subtypes and EF-subtypes in ADHD. Pfiffner et al. (24) argue that compared to ADHD-C, the primary inattentive subtype (ADHD-I) is characterized by more severe problems related to alertness/orientation, cognitive tempo, and processing speed, but on the other hand, shows fewer issues concerning impulsive and externalizing behavior.

The data in the present study seem to indicate a subscale difference where Inhibition (CANTAB) is negatively associated with Monitoring (BRIEF) in the combined group, and Working Memory (CANTAB) is negatively associated with Initiate and Monitoring (BRIEF) in the predominantly inattentive group. It is possible to argue that these results fit a cognitive model of ADHD-C as defined by difficulties in inhibition and ADHD-I as defined through difficulties related to the ability to initiate.

It is an interesting prospect to theorize how to best subdivide ADHD, and to test the best explanatory power, this is however outside the scope of the present study. It is also important to appreciate that these subscale differences are small compared to the overarching findings of the present study: Comparing EF cognitive test and EF rating scales, the results are primarily non-significant and muddled in children with ADHD. When comparing the present results to similar earlier findings, the few significant subscale correlations, that we found, seem to be random findings due to multiple comparisons.

In the analysis of the subgroup of participants with ADHD-I, the correlation between results from EF tests and EF rating scales becomes more significant and more opposing (negatively correlated).

Cross-methodological Biases

As argued by Toplak et al. (13), Ledochowski, Andrade & Toplak (16), and Kofler et al. (11), the lack of ecological validity involved in EF cognitive tests may partly explain the lack of expected cross-methodological correlation. One issue could be the way EF cognitive tests are conducted; in clinical settings with few competing stimuli and a single adult structuring the well-defined assignments, results from EF cognitive tests may lead to different outcomes as compared to how parents and teachers experience the child under everyday circumstances filled with many stimuli and expectations.

It is possible to expand the argument of ecological validity by including a critique of the way the results from different EF cognitive tests are combined to form a composite measure. In this way, the combined test results from isolated tasks, each measuring isolated cognitive abilities, are added together and treated as if this sum is representative of a higher-order EF cognitive function. However, it is possible to argue that this is not representative of how EF functions are applied in ecological settings. Rather, in the everyday life of problem-solving, many different cognitive and behavioral abilities are used simultaneously. The ability to drive a car, for example, demands the combined and simultaneous application of working memory, visuospatial thinking, long-term and procedural memory, inhibition, cognitive flexibility/shift, and more.

Besides the issue of discrepancy in ecological validity, the parents as sole informants on the rating scales in our study may also play a role. Teacher ratings represent an important source of information from an important context often defined by the higher demands and behavioral expectations of the school setting (25, 26). A study by Schneider, Ryan & Mahone (22), on the other hand, found that parents were prone to rate children as having more symptoms than teachers.

In the present study, the parents' responses to the rating scales may have been influenced by their knowledge of the research project they were participating in. The parents knew that the aim of the study (17) was to target organizational difficulties and that a lack of EF difficulties would lead to an exclusion from the study, and hence from the possibility of receiving a new type of treatment for their child.

In a study conducted by Tamm & Peugh (25) with 243 preschool children, the results indicated a small to moderate correlation between teacher ratings and EF test results, and more importantly, the strongest correlation was found between the Child Behavior Rating Scale's (CBRS) Behavioral Regulation subscale and the EF test results.

Tamm & Peugh's (25) results, as well as the results of the present study, connect to Soto et al. (15), and Kofler et al. (11) findings. That is, when comparing EF tests and rating scales, the results seem to suggest that EF tests and rating scales do in fact measure different underlying constructs. This perspective changes the narrative of the discussion. Perhaps the difference between tests and rating scales when measuring EF is not only a matter of ecological validity or rater bias, but is also related to the possibility that tests and rating scales measure different underlying constructs and expressions of EF (cognition and behavior respectively). Similarly, the most significant positive correlation achieved in the present study is between similar, but different constructs pertaining to cognition (Planning) and behavior (Organized Actions). This viewpoint demands new ways of understanding this complex relationship, new conceptualizations, and new ways of investigation.

Limitations

The study only includes parents as raters on BRIEF and COSS. We used a high score on BRIEF Organize to select our group of children. By including children with high scores on BRIEF, we could be limiting the distribution of scores and reducing the chances to detect significant correlations between BRIEF and CANTAB. The CANTAB test for working memory only concerns the non-verbal working memory. In the CANTAB test for planning, the item chosen measures correct answers on the first try, alternate items measuring errors might yield different results. The analysis did not include a specific comparison between ADHD-I and ADHD-C subtypes.

Conclusion

Comparing EF cognitive tests and EF rating scales, the results are primarily non-significant and unclear for the entire group of children with ADHD and executive dysfunctions. Results concerning the subgroup of participants with a primary inattentive subtype indicate that the correlation between results from EF tests and EF rating scales are more significant and more negatively associated. Cross-methodological biases such as differences in ecological validity, respondent bias, and differences in the underlying constructs measured by tests and rating scales can play a vital role in explaining the present findings. Further research in these cross-methodological biases is important to better understand the complex nature of EF as both cognition and behavior, how EF is examined in the clinic, and in order to understand and research the potential ways EF is expressed in subgroups of children and adolescents with ADHD.

Clinical implications

With the methods of assessing EF in ADHD available, the expectation should not be to find the same results when comparing data from tests and rating scales. It is relevant to interpret rating scales and cognitive tests as measuring different underlying constructs (behavioral and cognitive aspects).

Research implications

There is a need for research that takes a novel approach to: a) EF tests with a higher level of ecological validity, b) EF rating scales that are designed to capture EF behavior more so than EF cognition, and c) a research design that incorporates these considerations in a cross-methodological setup.

Funding
This study was funded by the Psychiatric Research Foundation of the Region of Southern Denmark and the Jascha Foundation.

Conflicts of Interest
The authors declare no conflict of interest.

References

1. Diamond A. Executive functions. Annu Rev Psychol. 2013;64:135-68.

2. Whittingham K, Bodimeade HL, Lloyd O, Boyd RN. Everyday psychological functioning in children with unilateral cerebral palsy: does executive functioning play a role? Dev Med Child Neurol. 2014;56(6):572-9.

3. Krivitzky L, Bosenbark DD, Ichord R, Jastrzab L, Billinghurst L. Brief report: Relationship between performance testing and parent report of attention and executive functioning profiles in children following perinatal arterial ischemic stroke. Child Neuropsychol. 2019;25(8):1116-24.

4. McLean RL, Johnson Harrison A, Zimak E, Joseph RM, Morrow EM. Executive function in probands with autism with average IQ and their unaffected first-degree relatives. J Am Acad Child Adolesc Psychiatry. 2014;53(9):1001-9.

5. de Vries M, de Ruiter MA, Oostrom KJ, Schouten-Van Meeteren AYN, Maurice-Stam H, Oosterlaan J, et al. The association between the behavior rating inventory of executive functioning and cognitive testing in children diagnosed with a brain tumor. Child Neuropsychol. 2018;24(6):844-58.

6. Krieger V, Amador-Campos JA. Assessment of executive function in ADHD adolescents: contribution of performance tests and rating scales. Child Neuropsychol. 2018;24(8):1063-87.

7. Claesdotter E, Cervin M, Akerlund S, Rastam M, Lindvall M. The effects of ADHD on cognitive performance. Nord J Psychiatry. 2018;72(3):158-63.

8. Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJ. Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? Biol Psychiatry. 2005;57(11):1224-30.

9. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol Psychiatry. 2005;57(11):1336-46.

10. Lambek R, Tannock R, Dalsgaard S, Trillingsgaard A, Damm D, Thomsen PH. Executive dysfunction in school-age children with ADHD. J Atten Disord. 2011;15(8):646-55.

11. Kofler MJ, Sarver DE, Harmon SL, Moltisanti A, Aduen PA, Soto EF, et al. Working memory and organizational skills problems in ADHD. J Child Psychol Psychiatry. 2018;59(1):57-67.

12. Wahlstedt C, Thorell LB, Bohlin G. Heterogeneity in ADHD: neuropsychological pathways, comorbidity and symptom domains. J Abnorm Child Psychol. 2009;37(4):551-64.

13. Toplak ME, Bucciarelli SM, Jain U, Tannock R. Executive functions: performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). Child Neuropsychol. 2009;15(1):53-72.

14. Toplak ME, West RF, Stanovich KE. Practitioner review: do performance-based measures and ratings of executive function assess the same construct? J Child Psychol Psychiatry. 2013;54(2):131-43.

15. Soto EF, Kofler MJ, Singh LJ, Wells EL, Irwin LN, Groves NB, et al. Executive functioning rating scales: Ecologically valid or construct invalid? Neuropsychology. 2020;34(6):605-19.

16. Ledochowski J, Andrade BF, Toplak ME. A novel unstructured performance-based task of executive function in children with attention-deficit/hyperactivity disorder. J Clin Exp Neuropsychol. 2019;41(5):445-59.

17. Bikic A, Dalsgaard S, Olsen KD, Sukhodolsky DG. Organizational skills training for children with ADHD: study protocol for a randomized, controlled trial. Trials. 2021;22(1):752.

18. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry. 1997;36(7):980-8.

19. Reynolds RK, C. Intellectual Assessment Scales and Reynolds Intellectual Screeining Test Professional Manual. Psychological Assessment Resources. 2003.: Psychological Assessment Resources; 2003.

20. Gioia GA, Isquith PK, Retzlaff PD, Espy KA. Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. Child Neuropsychol. 2002;8(4):249-57.

21. Gioia GA, Isquith PK, Guy SC, Kenworthy L.. Behavior Rating Inventory of Executive Function Child Neuropsychology. 2000;6(3):235-8.

22. Schneider H, Ryan M, Mahone EM. Parent versus teacher ratings on the BRIEF-preschool version in children with and without ADHD. Child Neuropsychol. 2020;26(1):113-28.

23. Abikoff H, Gallagher, R. . Children's Organizational Skills Scales manual. North Tonawanda, NY Multi-Health Systems; 2009.

24. Pfiffner LJ, Yee Mikami A, Huang-Pollock C, Easterlin B, Zalecki C, McBurnett K. A randomized, controlled trial of integrated home-school behavioral treatment for ADHD, predominantly inattentive type. J Am Acad Child Adolesc Psychiatry. 2007;46(8):1041-50.

25. Tamm L, Peugh J. Concordance of teacher-rated and performance-based measures of executive functioning in preschoolers. Child Neuropsychol. 2019;25(3):410-24.

26. Dekker MC, Ziermans TB, Spruijt AM, Swaab H. Cognitive, Parent and Teacher Rating Measures of Executive Functioning: Shared and Unique Influences on School Achievement. Front Psychol. 2017;8:48.