Journal of Psychopathology and Behavioral Assessment
https://doi.org/10.1007/s10862-022-09967-7

Consistently Inconsistent Working Memory Performance Among Children with ADHD: Evidence of Response Accuracy Variability (RAV)

Lauren M. Friedman[1] · **Mark D. Rapport**[2] · **Gabrielle Fabrikant-Abzug**[1]

Accepted: 15 March 2022

Abstract

Heterogeneity in cognitive performance, once regarded as noise, is now considered a causal mechanism or core deficit of ADHD and its related symptoms in most etiological models of the disorder. Previous research on cognitive performance variability has documented increased heterogeneity in response latencies using reaction time data. In contrast, variability in response accuracy remains understudied. The present study is the first to examine Response Accuracy Variability (RAV) among children with ADHD. Children with ADHD ($N=54$) and typically developing children ($N=50$) completed phonological working memory tasks with four set size conditions. RAV was calculated for each set size using the adjusted coefficient of variation (ζ). Results from a mixed model ANOVA indicated that children with ADHD evinced significantly greater variation in working memory performance relative to typically developing children when engaged in tasks within their cognitive capacity (i.e., set sizes 3 and 4), whereas all children exhibit similar, high levels of variability on tasks that exceeded their cognitive capacity (i.e., set sizes 5 and 6). Findings are aligned with the extant literature in documenting *consistently inconsistent* cognitive performance among children ADHD.

Keywords Attention-Deficit/Hyperactivity Disorder (ADHD) · Executive Functions · Working Memory · Heterogeneity

Attention Deficit/Hyperactivity Disorder (ADHD) is a commonly occurring neurodevelopmental disorder that affects approximately 7% of school-aged children (Thomas et al., 2015). Although a clear, consistent clinical picture entailing impairing inattention, hyperactivity, and impulsivity has been described for over two centuries (Lange et al., 2010), substantial evidence indicates that the ADHD phenotype is characterized by heterogeneity in symptom expression (Campez et al., 2020; Kofler et al., 2016; Orban et al., 2018), social and academic functioning (DuPaul et al., 2016; Kofler et al., 2017), and cognitive abilities (Mostert et al., 2018; Raiker et al., 2019; Wåhlstedt et al., 2009). As neurocognitive deficits are increasingly implicated in etiological models of the disorder, variability in cognitive performance has garnered significant interest in recent years. Once treated as a nuisance variable, or *noise*, heterogeneity in cognitive performance is considered a causal mechanism/core deficit of ADHD-related symptoms (Castellanos et al., 2006; Halperin & Schulz, 2006; Russell et al., 2006; Sonuga-Barke & Castellanos, 2007) or correlated outcome of underlying factors (Barkley, 1997; Leth-Steensen et al., 2000; Rapport et al., 2008; Sonuga-Barke et al., 2010) in most contemporary theoretical models of the disorder. Anecdotal and empirical evidence also documents situational and temporal variability in cognitive abilities (Orban et al., 2018; Roberts et al., 2015). For example, parents and clinicians often report that children with ADHD are able to complete a math problem in one moment (e.g., while completing a math homework worksheet or standardized achievement measure) but fail to correctly answer a nearly identical item mere minutes later. This has led to observations that children with ADHD are *consistently inconsistent* (Karalunas, 2010; Rapport, 1994). In sum, convincing empirical, etiological, and anecdotal data indicates that cognitive performance variability is a prominent feature of ADHD.

Variability in cognitive performance has traditionally relied on reaction time (RT) data to quantify moment-to-moment variation in cognitive performance. RT variability reflects participants' trial-by-trial speed of responding

✉ Lauren M. Friedman
Lauren.Friedman@asu.edu

1 Department of Psychology, Arizona State University, 900 S. McAllister Ave, Tempe, AZ 85281, USA

2 Department of Psychology, University of Central Florida, 4000 Central Florida Blvd, Orlando, FL 32816, USA

Published online: 02 April 2022

during laboratory tasks and examines RT dispersion as a means of operationalizing intraindividual variability in cognitive performance. RT variability has been studied extensively in children with ADHD, and extant meta-analytic evidence suggests that children with ADHD show greater RT variability relative to typically developing peers (Hedges' g = 0.76) and clinical control children (Hedges' g = 0.25; Kofler et al., 2013). The observed variability is often attributed to a subset of abnormally slow responses not observed in typically developing children, and appears consistent even after controlling for mean RT (Galloway-Long & Huang-Pollock, 2018; Kofler et al., 2013).

If cognitive variability is a prominent characteristic of ADHD, it is likely that processes most proximal to the etiological cognitive mechanisms of the disorder will show the largest performance variation. Yet, simple and choice RT tasks, which are oft used measures to index variability, require minimal executive resources and predominantly rely on recognition and signal detection rather than higher-order cognitive abilities for successful execution. Evidence for ADHD-related performance deficits on simple/choice RT tasks is mixed as children with ADHD display similar (Kofler et al., 2013) or faster (Raiker et al., 2019) mean RTs relative to peers. Behavioral inhibition tasks are also commonly utilized to assess RT variability. Although many children with ADHD show deficits in inhibitory control and behavioral disinhibition is often cited as a core deficit of ADHD (Barkley, 1997, 2015), accumulating evidence suggests that ADHD-related behavioral inhibition deficits are either weakly or unrelated to core behavioral symptoms (Brocki et al., 2010; Kuntsi et al., 2001; Nigg, 1999; Solanto et al., 2001) and better explained by basic attentional or working memory processes (Alderson et al., 2007, 2017; Lijffijt et al., 2005).

Examining working memory performance variability might be more appropriate than response speed or inhibitory control given evidence demonstrating robust working memory deficits among children with ADHD (see Rapport et al., 2013, for a review). Experimental and meta-analytic studies find large-magnitude differences in working memory performance accuracy (i.e., proportion of correct responses) among children with ADHD relative to typically developing children ($d = 2.01$–2.76, Kasper et al., 2012; Rapport et al., 2008). In addition, children with ADHD are more likely to evince working memory deficits relative to other areas of executive functioning such as behavioral inhibition or set shifting (Kofler et al., 2019). Working memory is also strongly related to ADHD core symptoms of inattention (Kofler et al., 2010; Orban et al., 2018), hyperactivity (Rapport et al., 2009; Sarver et al., 2015) and impulsivity (Raiker et al., 2012), as well as important functional outcomes such as academic underachievement (Calub et al., 2019; Eckrich et al., 2019; Friedman et al., 2017, 2018a), poor social skills (Kofler et al., 2011), and activities of daily living (Irwin et al., 2021). Given that children with ADHD exhibit large magnitude working memory deficits that are functionally related to core and secondary symptoms, children with ADHD may similarly show increased variability in working memory performance accuracy when examined trial-by-trial. Yet, no study to date has examined accuracy variability in working memory performance.

Several studies have examined working memory performance variability by examining RT distributions (i.e., response latencies) during working memory updating tasks, such as *N*-back tasks or the Paced Auditory Serial Addition Test (PASAT). Increased intraindividual RT variability was observed among children with ADHD relative to neurotypical peers (Buzy et al., 2009; Epstein et al., 2011; Fassbender et al., 2009; Karatekin, 2004), and this effect was largest for tasks with greater cognitive load (i.e., 1-back vs 0-back; Klein et al., 2006). While RT variability during working memory tasks is an important indicator of consistency in response speed, it does not index *accuracy* variability. Response Accuracy Variability (RAV), conversely, measures trial-by-trial responses to determine whether the proportion of correct responses (e.g., stimuli recalled correctly per trial) varies throughout the task. This latter metric is likely a more informative indicator of intraindividual working memory performance variability given that (a) extant literature documenting working memory deficits among children with ADHD primarily utilizes accuracy indices rather than reaction times, and (b) RAV is likely related to higher-order cognitive abilities, such as cognitive and attentional control (Fassbender et al., 2009). To our knowledge, no study to date has examined whether children with ADHD evince increased working memory RAV relative to peers. The present investigation examines this possibility.

While several metrics for assessing RT variability have been identified and validated, these indices are ill-suited for use with accuracy data. For example, a common approach for quantifying RT variability is to examine the standard deviation (SD) of reaction times across trials. However, SD is inappropriate for assessing variability in accuracy data. SD increases linearly with mean performance, and determining whether large SDs reflect lower mean accuracy or de facto increases in response variability is problematic. Using SD to analyze accuracy data also artificially restricts variability when accuracy performance is very high or low, preventing valid examination of RAV (see Golay et al., 2013, for a review).

Another approach commonly used to assess intraindividual variability is the coefficient of variation (CV). CV is measured by dividing the SD by the mean to obtain a dimensionless number, usually expressed as a percent or proportion, that is theoretically unrelated to the mean. For this reason, CV is preferred over SD for RT data, as it is largely unaffected by mean performance. However, accuracy data behaves differently and shares large variance with mean performance when analyzed using the CV, primarily due to differences in the range of potential values in RT data relative to accuracy.[1] Therefore, the use of CV metrics to index RAV "could be a very redundant and potentially misleading indicator" of accuracy variability (p. 12, Golay et al., 2013).

Recent ex-gaussian approaches for examining RT data are also ill-suited for accuracy data. This approach separates children's RT distributions into exponential (*ex*) and normal (*gaussian*) functions to discern whether observed RT variability reflects a proportion of abnormally slowed responses. The positively skewed RTs within the exponential component reflect slowed responses, and children with ADHD show an increased proportion of responses within this distribution relative to typically developing children (Kofler et al., 2013; Tamm et al., 2012). However, the ex-gaussian approach requires values to vary freely on the upper limit (i.e., values must be infinitely large) to accurately model and dissociate the exponential function. That is, skew in accuracy is not evident if values are artificially truncated at the maximum (or minimum) possible value (e.g., 5 out of 5 correct responses). Consequently, ex-gaussian analyses are ill-suited for examining accuracy data.

A novel approach for measuring RAV without the limitations noted above uses the adjusted coefficient of variation (ζ). This metric is calculated by dividing the observed variability by the maximum possible variability at a given set size (see Golay et al., 2013, for formula derivation). Similar to the traditional CV, ζ ranges from 0 to 1 but is unrelated to mean performance when data values are restricted on both ends of the continuum, as is the case with accuracy data. While ζ is sensitive to task difficulty when difficulty is either very high or low, prudent selection of item set size renders ζ an accurate and valid indicator of intraindividual variability in performance accuracy (Golay et al., 2013; Mestdagh et al., 2018).

The present study is the first to examine ADHD-related cognitive performance variability by measuring accuracy inconsistency. A novel approach for studying accuracy variability, the adjusted coefficient of variation (ζ), was selected to minimize the potential confounds of relying on RT data to characterize cognitive performance variability and circumvents the shortcomings inherent to traditional variability metrics. Children with ADHD were hypothesized to exhibit increased trial-by-trial variability in working memory response accuracy, as measured by ζ, relative to typically developing peers. We also hypothesized that significant relations among RAV and ecologically valid outcomes such as academic achievement, home/school functioning, and social problems would be evident based on the expectation of finding significant between group differences in accuracy variability. If supported, the results will serve to further elucidate the potential etiological mechanisms of ADHD and help inform the design of theory-driven interventions for the disorder.

Methods

Participants

The current study is a secondary data analysis of a larger experimental study to understand the neurocognitive contributors to ADHD-related symptoms and impairment (Kofler et al., 2008, 2010; Raiker et al., 2012; Rapport et al., 2009). The sample comprised 104 boys aged 8 to 12 years ($M = 9.29$, $SD = 1.30$) recruited or referred to a children's learning clinic through community resources (e.g., local pediatricians, community mental health clinics, and school mental health professionals, as well as self-referrals). Sample race and ethnicity included 71 Caucasian (68.3%), 19 Hispanic/Latinx (18.3%), 5 African American (4.8%), and 8 bi- or multi-racial youth (7.7%), and 1 (1%) child who indicated 'other' racial or ethnic identity. All procedures were approved by the University's Institutional Review Board, and all parents and children provided their informed consent and assent, respectively, prior to study participation. Two groups of children participated in the study: children with ADHD Combined Presentation ($n = 54$), and typically developing children ($n = 50$) without a psychological disorder. Children with a history of (a) gross neurological, sensory, or motor impairment by parent report, (b) history of a seizure disorder by parent report, (c) psychosis, or (d) Full Scale IQ score < 85 were excluded.

Group Assignment

All children and their parents participated in a detailed, semi-structured clinical interview using all modules of the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS). The K-SADS assesses onset, course, duration, severity, and impairment of current and past episodes of psychopathology in children and adolescents based on DSM criteria. Its psychometric properties are well established, including interrater agreement of .93 to 1.00, test–retest reliability of 0.63 to 1.00, and concurrent (criterion) validity between the K-SADS and psychometrically established parent rating scales (Kaufman et al., 1996).

[1] RT data are restricted on the lower end to be greater than zero but could vary within an infinitely large set of possible values on the upper end. Accuracy data exist on a finite continuum bounded on the lower end by 0 and on the upper end by the number of stimuli presented (e.g., 5 out of 5 stimuli correct).

Fifty-four boys meeting the following criteria were included in the ADHD group: (1) an independent diagnosis by the directing clinical psychologist using DSM-5 criteria for ADHD, combined presentation, based on K-SADS interview with parent and child; (2) parent ratings of at least 2 SDs above the mean on the Attention-Deficit/Hyperactivity Problems DSM-Oriented scale of the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2003), or exceeding the criterion score for the parent version of the ADHD-Combined subtype subscale of the Child Symptom Inventory-4: Parent Checklist (CSI-P; Gadow & Sprafkin, 2002); and (3) teacher ratings of at least 2 SDs above the mean on the Attention-Deficit/Hyperactivity Problems DSM-Oriented scale of the Teacher Report Form (TRF; Rescorla & Achenbach, 2004), or exceeding the criterion score for the teacher version of the ADHD-Combined subtype subscale of the Child Symptom Inventory-4: Teacher Checklist (CSI-T; Gadow & Sprafkin, 2002). The CBCL, TRF, and CSI are among the most widely used behavior rating scales for assessing psychopathology in children. Their psychometric properties are well established (Rapport et al., 2008). Twenty (37%) of the children with ADHD were on a psychostimulant regimen for treatment of their ADHD symptoms (24-h washout period prior to each testing session), and five (9.3%) met diagnostic criteria for Oppositional-Defiant Disorder (ODD).

Fifty boys met the following criteria and were included in the typically developing group: (1) no evidence of any clinical disorder based on parent and child K-SADS interview; (2) normal developmental history by parental report; (3) ratings within 1.5 SDs of the mean on all CBCL and TRF scales; and (4) parent and teacher ratings within the non-clinical range on all CSI subscales.[2]

Measures

Phonological Working Memory. The Phonological Working Memory task used in the current study is described extensively in Rapport et al. (2008). Briefly, the task is similar to the Letter-Number Sequencing subtest on the WISC-V (Wechsler, 2014). Children were presented a series of numbers and a capital letter one at a time on a computer monitor and instructed to recall the numbers in order from smallest to largest and place the letter last (e.g., 8 L 9 3 is correctly recalled as 3 8 9 L). Each number and letter (4 cm height) appeared on the screen for 800 ms, followed by a 200 ms interstimulus interval. Four phonological conditions (i.e., set sizes 3, 4, 5, and 6) were administered. The four working memory set size conditions each contained 24 unique trials of the same stimulus set size and were counterbalanced across four testing sessions to control for order effects and potential proactive interference effects across set size conditions (i.e., children completed 24 trials at a single set size during each of four testing sessions). The letter never appeared in the first or last position of the sequence to minimize potential primacy and recency effects, and trials were counterbalanced to ensure that letters appeared an equal number of times in the other serial positions (i.e., position 2, 3, 4, or 5). All children completed five practice trials prior to each administration and achieved the minimum of 80% accuracy on training trials. The working memory task has high internal consistency ($\alpha = .81$) in the current sample and the expected level of external validity ($r = .50$ to .66) with WISC-III and -IV Digit Span raw scores (Raiker et al., 2012).

Functional Outcomes. The following functional outcomes were used to assess the predictive validity of RAV in exploratory models. *Overall impairment at home* was assessed using the symptom severity score of the Home Situations Questionnaire (HSQ). The HSQ is a 12 item, parent-reported rating scale that assesses children's overall impairment at home across situations on a 9-point scale ranging from 1 (mild) to 9 (severe). *Parent-rated overall functioning at school* was assessed using age- and gender-corrected T-scores from the School Competence Scale scores from the Child-Behavior Checklist (CBCL). The School Competence Scale assesses parent-reported performance at school across four academic subjects (Reading, Math, History, and Science) on a four-point scale ranging from 1 (failing) to 4 (above average). *Teacher-rated overall functioning at school* was assessed using age- and gender- corrected T-scores from the Academic Competence and the Working Hard, Behaving, and Learning Adaptive Functioning subscales of the Teacher Report Form (TRF). *Overall social impairment* was assessed using age- and gender-corrected T-scores from the Social Problems subscale of the CBCL. The CBCL Social Problems subscale assesses several domains of social impairment including peer rejection, social interaction style, and perceived impact of social impairments. Endorsements are provided on a 3-point scale ranging from 0 (not true) to 2 (very true or often true). Finally, *overall academic achievement* was assessed using the Kaufman Test of Educational Achievement (KTEA)-first (normative update, Kaufman, 1997), second (Kaufman, 2004), or third (Kaufman, 2014) editions ($r = .74$ to .93 between versions for all subtests; Kaufman, 2004, 2014).

[2] Scores for one TD child exceeded 1.5 *SD*s on one of the two parents' but not teachers' rating scales. Parent interview revealed no significant ADHD symptoms or symptoms associated with other clinical disorders for the child. Six children with ADHD had subthreshold scores on teacher-rated hyperactivity/impulsivity. Follow-up clinical interviews, however, indicated the subthreshold symptoms were attributable to substantial psychostimulant effects while they were rated, and that all children demonstrated a history of significant, persistent levels of hyperactivity/impulsivity both at home and at school.

Table 1 Sample demographic characteristics

Variable	ADHD		Typically Developing			
	M	*SD*	M	*SD*	*t*	Cohen's *d*
Age	9.06	1.29	9.54	1.28	1.92	-0.37
FSIQ	103.87	12.10	107.56	12.03	1.56	-0.31
$FSIQ_{res}$	-.03	.99	.02	1.01	.25	-0.05
SES	49.35	10.06	53.36	9.66	2.03*	-0.41
CBCL ADHD DSM-Oriented	72.06	6.96	53.94	6.02	-14.01***	2.85
TRF ADHD DSM-Oriented	67.22	6.98	53.34	4.76	-11.46***	2.32
CSI-P: ADHD, Combined	77.69	9.66	49.73	11.18	-13.41***	2.68
CSI-T: ADHD, Combined	69.49	9.37	48.69	8.10	-11.65***	2.37

ADHD attention-deficit/hyperactivity disorder, *CBCL* Child Behavior Checklist, *CSI-P* Child Symptom Inventory: Parent severity *T*-scores, *CSI-T* Child Symptom Inventory: Teacher severity *T*-scores, *FSIQ* Full Scale Intelligence Quotient, $FSIQ_{res}$ Full Scale Intelligence Quotient with working memory removed, *SES* socioeconomic status, *TRF* Teacher Report Form

$^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$

The change-over to newer versions was due to their release during the study and to provide parents the most up-to-date educational evaluation possible.[3] Age-corrected standard scores from the Battery Composite (KTEA-I), Comprehensive Achievement Composite (KTEA-II), or Academic Skills Battery Composite (KTEA-III) were used to assess overall academic achievement.

Data Analytic Plan

All statistical analyses were performed using IBM Corp. (2019). Preliminary analyses involved investigation of missing data and examination of demographic characteristics for potential between group differences (see Table 1). Primary analyses involved mixed model analyses of variance (ANOVAs) examining within (Working Memory Set Size) and between (Typically Developing; ADHD) group effects. Separate models were run to examine Accuracy and Response Accuracy Variability (RAV). Consistent with best practice recommendations, stimuli correct per trial were used (Conway et al., 2005; Kasper et al., 2012; Wells et al., 2018), and separate scores at each working memory set size were derived. Performance Accuracy reflected the intraindividual mean stimuli correct per trial at each set size. RAV was calculated for each set size using the adjusted coefficient of variation (ζ), as described by Golay and colleagues (2013). Briefly, ζ provides an index of the ratio between the observed intraindividual variability and the maximum possible variability at any given level of task difficulty. ζ is derived using the following equation:

$$\zeta = \frac{iSD}{\sqrt{\frac{n}{n-1}}\sqrt{iM\,D - iM^2}}$$

where *iSD* is the intraindividual standard deviation, *n* is the number of trials, *iM* is the intraindividual mean stimuli correct per trial, and *D* is the range of possible accuracy values (e.g., 4 during a set size 4 task). ζ ranges from 0 to 1 and is not related linearly to mean performance when applied to accuracy data–a key improvement over extant performance variability metrics. To reduce the likelihood of Type I error, post-hoc analyses were performed using the Benjamini–Hochberg False Discovery Rate (FDR; Benjamini & Hochberg, 1995) applied within domain. The FDR is associated with lower rates of familywise error relative to other approaches (e.g., Bonferroni correction). For all pairwise comparisons, Cohen's *d* effect size metrics are provided (0.2 = small, 0.5 = moderate, 0.8 = large).

Analyses were initially completed without covariates. We then performed exploratory ANCOVAs to examine the following possible covariates: participant age, race, medication status, and socioeconomic status (Hollingshead, 1975). Inclusion of the identified covariates did not change the pattern or interpretation of results. Therefore, simple models without covariates are presented. We did not include IQ as a covariate, consistent with best practice recommendations (Dennis et al., 2009; Miller & Chapman, 2001). That is, working memory shares significant variance (r = .68 to .79) with Full Scale IQ, and removal of variance attributable to FSIQ would remove important variance in working memory from working memory–the key variable of interest in the present study. Consistent with past studies (Friedman et al., 2017, 2018b; Rapport et al., 2008), we removed reliable variance associated with working memory from FSIQ, and then examined between group differences in FSIQ without

[3] KTEA version was examined as a potential covariate in relevant models but was nonsignificant.

Table 2 Working Memory Accuracy Analyses

	Set Size 3 M (SD)	Set Size 4 M (SD)	Set Size 5 M (SD)	Set Size 6 M (SD)	Group Composite M (SE)	Set Size F
ADHD	2.6 (.39)	3.08 (.59)	2.85 (.99)	2.48 (1.06)	2.76 (.07)	32.299***
	86.7%	77.0%	57.0%	41.3%		
TD	2.86 (.17)	3.65 (.39)	3.98 (.61)	3.75 (1.00)	3.56 (.08)	80.046***
	95.3%	91.3%	79.6%	62.5%		
Set Size Composite	2.73 (.33)	3.36 (.58)	3.40 (1.00)	3.09 (1.21)		
	91.0%	84.0%	68.0%	51.5%		
Group F	19.126***	33.329***	47.638***	39.501***		
Group Contrasts	TD > ADHD	TD > ADHD	TD > ADHD	TD > ADHD		
	(d = -0.86)	(d = -1.52)	(d = -1.37)	(d = -1.23)		

The reported percentage reflects the percent of stimuli recalled correctly per trial

ADHD Attention-Deficit/Hyperactivity Disorder, *TD* Typically Developing

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

the influence of working memory. Results revealed that between-group differences in the residual FSIQ score were not significant ($p = .81$). As a result, simple model results without covariates are presented.

Results

Preliminary Analyses. All independent and dependent variables were screened for multivariate outliers using Mahalanobis distance tests (p < .001), and none were identified. Missing data represented 0.2% of all possible data points due to non-administration of the Working Memory Set Size 5 condition for one child with ADHD. Based on recommendations (Tabachnick & Fidell, 2007), the missing data point was replaced with the ADHD group mean. The inclusion or exclusion of this case did not change the pattern or interpretation of results. As expected, parent and teacher rating scales were significantly higher for the ADHD group relative to the TD group (see Table 1). Diagnostic groups did not differ on age ($p = .06$), race ($p = .50$), or FSIQ ($p = .12$). A slight difference in SES ($p = .05$) was found, but its inclusion as a covariate did not change the pattern or interpretation of results. Therefore, simple models without covariates are presented.

Working Memory Accuracy. As a first step and manipulation check, between group differences in Working Memory Accuracy were analyzed in a 2 (Diagnostic Status: TD vs ADHD) × 4 (Set Size: 3, 4, 5, 6) mixed model ANOVA. Means comparisons are shown in Table 2. As expected, a significant main effect of Diagnostic Status was observed, $F(1,102) = 56.651$, $p < .001$, indicating large magnitude ($\eta^2 = .36$) between-group differences on working memory accuracy. A significant main effect of Set Size, $F(3,306) = 33.472$, $p < .001$ and Diagnostic Status X Set Size interaction, $F(3,306) = 19.226$, $p < .001$, were also observed. Post-hoc analyses using the Benjamini–Hochberg FDR correction indicate that children with ADHD exhibited significantly reduced working memory accuracy relative to typically developing children across all set sizes ($d = -0.86$ to -1.52; all ps < .001; see Fig. 1a). However, the pattern of performance differed between the groups. Within subjects pairwise comparisons found that Typically Developing children showed a significant increase in the number of stimuli recalled correctly as Set Size increased from 3 to 4 ($p < 0.001$), and the number of stimuli recalled correctly per trial was maintained at a relatively consistent level ($M = 3.65$ to 3.98 stimuli correct per trial) as the Set Sizes increased from 4 to 5 ($p = .06$) and from 5 to 6 ($p = .44$). Conversely, children with ADHD showed a significant yet marginal increase in the number of stimuli recalled correctly from Set Sizes 3 to 4 ($p < .001$), and the number of stimuli correct per trial decreased steadily as cognitive load was increased from 4 to 5 ($p = .02$) and from 5 to 6 ($p = .001$).

Working Memory Response Accuracy Variability. Potential between-group differences in within-task working memory Response Accuracy Variability (RAV), as indexed by ζ, were examined subsequently. Means comparisons are shown in Table 3. A 2 (Diagnostic Status: TD vs ADHD) × 4 (Set Size: 3, 4, 5, 6) mixed model ANOVA revealed a significant main effect of Diagnostic Status, $F(1,102) = 12.749$, $p = 0.001$, indicating moderate to large magnitude ($\eta^2 = .11$) between-group differences on working memory RAV. A significant main effect of Set Size, $F(3,306) = 3.282$, $p = .021$, and Diagnostic Status X Set Size interaction, $F(3,306) = 6.209$, $p < .001$, was also observed. Post-hoc analyses using the Benjamini–Hochberg FDR correction indicate that significantly higher RAV was observed in children with ADHD relative to Typically Developing children at the two lowest

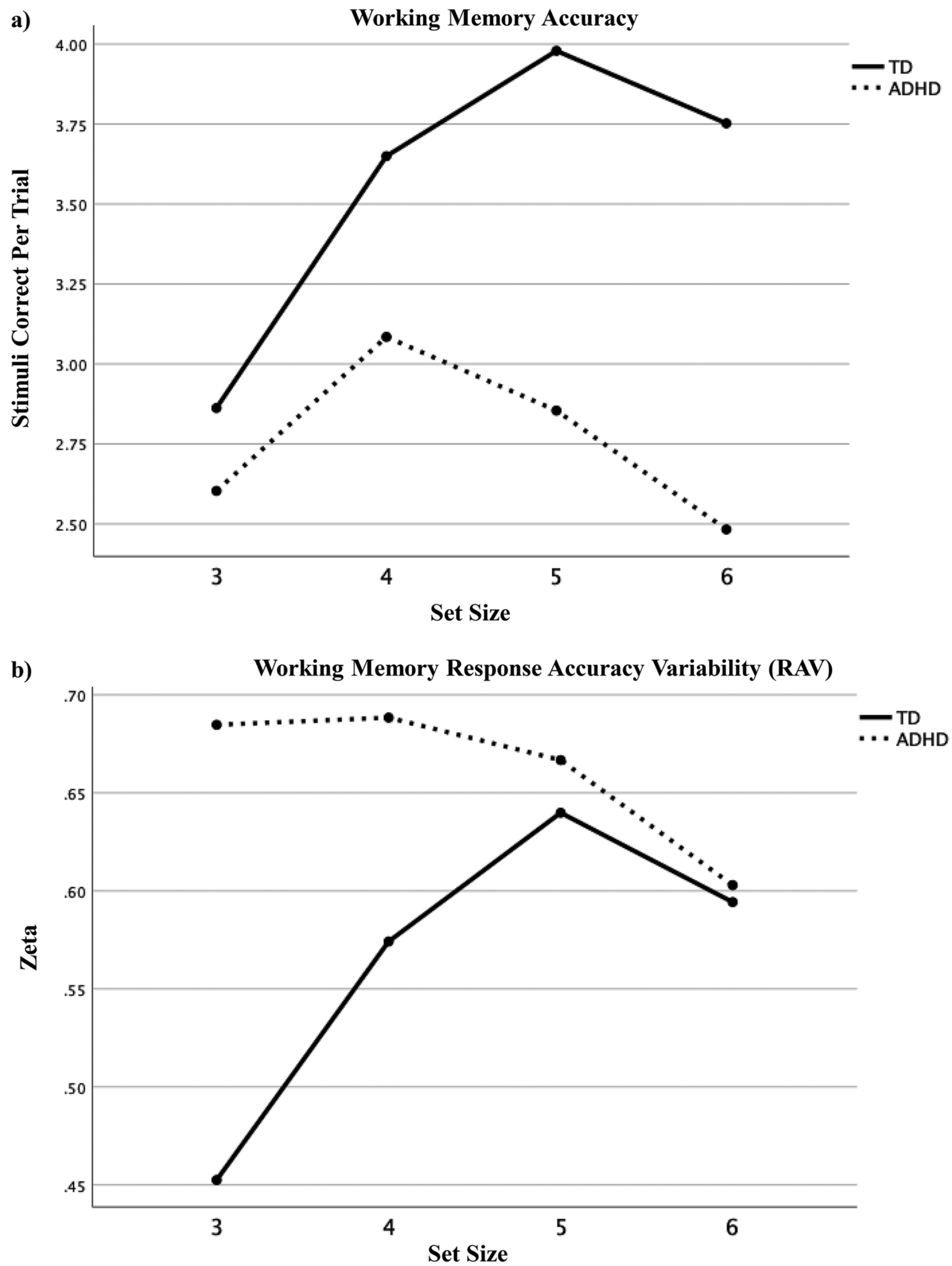

Fig. 1 Mean stimuli correct per trial (**a**) and Response Accuracy Variability (ζ) (**b**) as a function of set size and diagnostic group. *ADHD* Attention-Deficit/Hyperactivity Disorder, *TD* Typically Developing

set sizes (Set Size 3: $p=.001$, $d=0.66$; Set Size 4: $p=.007$; $d=0.57$) but not under the two highest set size conditions (Set Size 5: $p=.395$, $d=0.19$; Set Size 6: $p=.697$; $d=0.09$; see Fig. 1b). Within subjects pairwise comparisons revealed that children with ADHD showed similarly elevated rates of variability across all working memory load conditions, and no significant pairwise comparisons were evident as Set Size increased with one exception–children with ADHD showed a small magnitude decrease in variability as Set Size increased from 5 to 6 ($p=.005$, $d=0.37$). In contrast, Typically Developing children evinced significant increases in RAV as Set Size increased from 3 to 4 ($p=.019$) and from 4 to 5 ($p=.048$), followed by similar levels of variability as Set Size increased from 5 to 6 ($p=.06$). Collectively, results reveal that working memory performance among children with ADHD was characterized by high levels of RAV regardless of task difficulty, whereas Typically Developing children evinced reduced rates of within-task performance variability at lower set sizes but were indistinguishable from children with ADHD at higher set sizes.

Table 3 Working Memory Response Accuracy Variability (RAV) Analyses

	Set Size 3 M (SD)	Set Size 4 M (SD)	Set Size 5 M (SD)	Set Size 6 M (SD)	Group Composite M (SE)	Set Size F
ADHD	.68 (.31)	.69 (.18)	.67 (.17)	.60 (.10)	.66 (.02)	5.993**
TD	.45 (.38)	.57 (.24)	.64 (.15)	.59 (.12)	.56 (.02)	3.758*
Set Size Composite	.57 (.37)	.63 (.22)	.65 (.16)	.60 (.11)		
Group F	11.422**	7.669**	0.731	0.153		
Group Contrasts	TD < ADHD (d = 0.66)	TD < ADHD (d = 0.57)	TD = ADHD (d = 0.19)	TD = ADHD (d = .09)		

ADHD Attention-Deficit/Hyperactivity Disorder, *TD* Typically Developing

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

RAV and Relations with Ecologically Valid Outcomes. Given that neurocognitive variability is considered a core deficit (Russell et al., 2006; Sonuga-Barke & Castellanos, 2007) or correlated outcome (Rapport et al., 2008; Sonuga-Barke et al., 2010) of ADHD in most current etiological models of the disorder, a series of exploratory analyses were performed to examine the extent to which RAV is predictive of important functional outcomes, such as impairments in home and school functioning, social skills, and academic achievement, that are theorized to be secondary to deficits in cognitive abilities. A mean composite of RAV at Set Sizes 3 and 4 is used in the ensuing analyses[4] to facilitate valid assessment of accuracy variability.

Bias-corrected, bootstrapped correlation analyses using 5000 resamples with replacement revealed significant, moderate magnitude relations between RAV and functional outcomes across domains including parent-rated impairment at home ($r = .28$), parent-rated functioning at school ($r = -.27$), teacher-rated functioning at school ($rs = -.28$ to $-.43$), academic achievement ($r = -.32$), and social problems ($r = .28$), see Table 4. Regarding academic achievement, exploratory regression models were analyzed to determine whether RAV predicts achievement over and above Full Scale IQ–a well-documented and strong predictor of academic achievement (Calub et al., 2019). Bias-corrected, bootstrapped linear regression analyses[5] revealed that Full Scale IQ score showed large-magnitude relations to overall academic achievement, as expected ($R^2 = .42$, $\beta = .65$, $p < .001$). When RAV was added to the model, both Full Scale IQ ($\beta = .65$, $p < .001$) and RAV ($\beta = -.18$, $p = .049$) significantly predicted academic achievement and accounted for 44% of the variance in achievement scores ($R^2 = .444$), indicating that variability in working memory abilities makes significant contributions to academic achievement over and above IQ alone.

Discussion

A common observation and oft touted enigma conveyed by adults–that children with ADHD can complete a cognitive task accurately in one moment but fail to do so mere

Table 4 Bias corrected, bootstrapped correlations between Set Size 3 and 4 composite Response Accuracy Variability (RAV) and ecologically valid functional outcomes

	r (95% Confidence Interval)
HSQ Severity	0.28* (0.05, 0.48)
CBCL School Competence	-0.27* (-0.08, -0.46)
TRF Academic Performance	-0.28* (-0.13, -0.45)
TRF Working Hard	-0.30* (-0.08, -0.50)
TRF Behaving	-0.29* (-0.05, -0.52)
TRF Learning	-0.43*** (-0.25, -0.60)
CBCL Social Problems	0.28* (0.49, 0.48)
KTEA Overall Academic Achievement	-0.32* (-0.14, -0.48)

CBCL Child Behavior Checklist, *HSQ* Home Situations Questionnaire, *KTEA* Kaufman Test of Educational Achievement; *TRF* Teacher Report Form

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

[4] While ζ overcomes many of the limitations inherent to traditional variability metrics when applied to accuracy data (e.g., shares no linear relation to mean performance), a limitation associated with the statistic is that values may become unreliable when task difficulty is too high or low (see Golay et al., 2013). This phenomenon occurred under the two largest set size conditions (5 and 6), characterized by either a precipitous decline (ADHD) or no increase in cognitive performance (TD children) as the number of stimuli to be recalled exceeded four. Consequently, a mean composite score of RAV at set sizes 3 and 4 was used to best evaluate the outcomes.

[5] WISC and KTEA version was examined as a potential covariate but failed to reach significance; therefore, a simple model without covariates is presented.

seconds later–served as the impetus for the current study. It represents the first empirical investigation to focus on variability within the accuracy component of children's cognitive performance, rather than reaction time, and examines whether increased response accuracy variability (RAV) is present among children with ADHD.

Obtained results revealed that children with ADHD evince significantly greater trial-by-trial variation in working memory performance relative to typically developing children when engaged in tasks within their cognitive capacity (i.e., the ability to process 3 to 4 stimuli correctly during working memory trials). Moderate to large magnitude between group differences were evident on RAV metrics, and increased variability in performance accuracy was significantly related to important, ecologically valid outcomes such as academic achievement, social skills, and overall functioning in school and home settings. Results confirm model predictions that children with ADHD show increased variability in working memory accuracy, and that increased RAV is predictive of ADHD-related symptoms and impairments. It is interesting to note that significant between-group differences were not evident under working memory conditions that exceeded children's cognitive capacity. That is, children in both groups showed a similar and high level of working memory accuracy variability (indexed by ζ) under the two highest (5 and 6) stimulus set size conditions. This finding was expected based on extant literature indicating limited validity for the RAV metric when task demands exceed children's cognitive capacity and exponentially reflect an unbalanced ratio of errors to correct responses.

In addition, our study is the first to provide direct evidence linking inconsistency in working memory performance to functional impairments resulting from ADHD sequelae. RAV showed significant, moderate to large-magnitude relations to functioning at home and school, social interactions, and objectively measured academic achievement. Based on the well-documented linkages between ADHD-related working memory deficits and functional impairments within the reported in the literature (Eckrich et al., 2019; Friedman et al., 2017, 2018a; Kofler et al., 2011, 2017), we speculate that many, if not most, of these outcomes reflect inconsistent working memory functioning to a significant extent due to their reliance on an active and accurate working memory system.

One potential explanation for the increased RAV observed among children with ADHD is that the temporary lapses in intraindividual accuracy are secondary to the well-documented deficits in the executive attention network associated with the disorder (Fan et al., 2005), similar to proposed explanations for increased ADHD-related reaction time variability (Castellanos et al., 2005; Castellanos & Tannock, 2002; Gallo & Posner, 2016; Goos et al., 2009). The executive attention network comprises a series of interconnected cortical structures (lateral prefrontal and anterior cingulate cortices; Matsumoto & Tanaka, 2004) responsible for the regulation of cognitive processes that enable goal directed behaviors. Both substrates within the neural network are consistently identified as underdeveloped and/or underactive among children with ADHD (Bledsoe et al., 2013; Shaw et al., 2007) and show strong involvement while completing working memory tasks (Emch et al., 2019; Wager & Smith, 2003). Children with ADHD show increased momentary fluctuations in executive control relative to peers (Feige et al., 2013; Helps et al., 2011), and these variations may underlie the inconsistency in response accuracy observed among youth with ADHD in the present study. From a cognitive perspective, these fluctuations may suggest that typically developing children may be better able to direct attentional resources to maintaining information in WM, resulting in less variability in their performance (Martin et al., 2021). These propositions were not tested directly, and complementary neuroimaging and neurocognitive investigations are necessary to determine the extent to which increased RAV is associated with fluctuations in executive control network dysfunction.

Our findings complement extant evidence demonstrating that heterogeneity is a prominent feature of the disorder. Cross-situational variation has been well-studied, and children with ADHD exhibit higher rates of symptom expression in certain contexts but not others (e.g., classroom vs leisure activities, Orban et al., 2018; lunch/recess vs reading/math instruction, Porrino et al., 1983). Patterns of impairments are also varied. Some children evince deficits in academic, social, and/or family functioning, whereas others do not (Kofler et al., 2017). Between-task cognitive performance heterogeneity has also been reported, wherein the overwhelming majority of children with ADHD (89%) exhibited significant deficits in at least one of three core executive functions (working memory, behavioral inhibition, set shifting), but only 4% displayed significant performance decrements in all three areas (Kofler et al., 2019).

The present findings are also consistent with extant literature documenting increased within task heterogeneity as assessed by reaction time (RT) variability. Meta analytic evidence indicates that children with ADHD evince increased variability in response latencies regardless of the metric used (e.g., standard deviation, coefficient of variability, spectral power, ex-gaussian analyses; Kofler et al., 2013). Ex-gaussian analyses characterizing the pattern of RT variability reveal that children with ADHD evince a greater proportion of abnormally slowed responses (Kofler et al., 2013). As noted previously, these approaches are unsuitable for studying cognitive-based accuracy performance and resulted in our adopting the coefficient of variation (ζ) for this purpose. The results of the current study were nevertheless consistent with the more general premise and converging evidence that heterogeneity is a prominent characteristic of ADHD, and provides additional evidence regarding their *consistently inconsistent* cognitive performance (Rapport, 1994).

Despite significant methodological strengths (e.g., multi-method/multi-informant diagnostic assessment, well-validated working memory tasks, careful selection of task difficulty when assessing RAV), several limitations warrant discussion. The exclusive inclusion of boys within the current study reflects well-documented gender differences in reaction time variability (DeRonda et al., 2021), neurocognitive functioning (Bálint et al., 2009), neural structure (Dirlikov et al., 2015), and ADHD symptom presentation (Gaub & Carlson, 1997). While utilizing a narrow yet rigorously defined inclusion criteria has the benefit of improving internal validity, it also limits generalization to other populations. Future studies are needed to confirm the presence of increased accuracy variability among girls with ADHD, greater age-ranges of patients (e.g., younger children, adolescents, and adults), other ADHD presentations, and samples with greater diversity in socioeconomic status and racial/ethnic backgrounds. It is also necessary to assess whether increased intraindividual accuracy variability is evident in other working memory modalities (i.e., visuospatial working memory) and paradigms (e.g., updating, complex span), as well as other cognitive processes (e.g., short and long-term memory, delay discounting). Further, the present investigation examined RAV using a single type of working memory task. Future studies should examine RAV using multiple working memory indicators to remove task-specific variance and obtain a purer estimate of the construct (Conway et al., 2002, 2005; Engle et al., 1999).

Future studies are also needed to assess whether the increased rates of RAV identified in the present study have implications for understanding ADHD etiological factors. That is, the present study is the first to (a) identify the presence of increased accuracy variability among children with ADHD and (b) demonstrate that increased RAV is a robust predictor of important functional outcomes, including academic achievement. Because most theoretical models of ADHD consider neurocognitive variability to be a core deficit or important correlate of etiological factors, future studies are needed to confirm whether working memory accuracy variability has similar etiological implications. In addition, the present study assessed for the presence of increased RAV relative to typically developing children. While this is an important first step towards understanding the consistency of neurocognitive performance among children with ADHD, future studies should include other clinical comparison groups where neurocognitive deficits are implicated (e.g., learning disability, autism spectrum disorder, anxiety, and depression) to determine whether increased RAV is specific to ADHD-related sequelae or an indicator of psychopathology defined broadly.

The present findings have several noteworthy clinical implications. Currently, behavioral interventions, such as classroom management programs and behavioral parent training, are the predominant nonpharmacological interventions recognized as empirically supported treatments for ADHD (Wolraich et al., 2019). While behavioral interventions are considered to be 'gold standard' treatments for the disorder, 55–66% of children continue to evince significant symptoms and impairment following intervention (DuPaul et al., 2016; Swanson et al., 2001). The present findings may inform approaches to optimize behavioral interventions to improve treatment response. Specifically, results suggest that *consistent* cognitive and academic performance should be targeted during behavioral treatment rather than accuracy or productivity alone. Behavioral intervention strategies may include fixed or variable interval schedules of reinforcement, wherein positive contingencies are provided following *sustained* performance over a predetermined timeframe. Behavioral interventions should also target variables such as percent accuracy while completing home and school work to scaffold consistent performance on cognitively demanding tasks. Clinical trials employing these techniques have shown promise for reducing ADHD-related impairments (Hornstra et al., 2021; Kaminski et al., 2008; Rapport et al., 1980), and the present study provides an etiological rationale for incorporating time-based reinforcement schedules to reduce performance variability in children with ADHD.

As our understanding of the neurocognitive origins to ADHD grows, it is increasingly apparent that heterogeneity in symptom expression, functional impairments, and cognitive performance is the norm rather than the exception among children with ADHD. The present study adds to the growing literature that variability is an important data signal for understanding the origins of ADHD-related symptoms and impairments instead of noise that must be eliminated.

Funding None to declare.

Declarations

Ethical Approval Institutional Review Board approval from the University of Central Florida was obtained prior to and maintained throughout data collection. Informed consent/assent was obtained from all parent and child participants, respectively.

Conflict of Interest Lauren M. Friedman, Mark D. Rapport, and Gabrielle Fabrikant-Abzug declare that they have no conflict of interest.

Experiment Participants All procedures performed were in accordance with the ethical standards of the institution and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms & profiles.* University of Vermont, Research Center for Children, Youth, and Families.

Alderson, R. M., Patros, C. H. G., Tarle, S. J., Hudec, K. L., Kasper, L. J., & Lea, S. E. (2017). Working memory and behavioral inhibition in boys with ADHD: An experimental examination of competing models. *Child Neuropsychology, 23*(3), 255–272.

Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: A meta-analytic review of the stop-signal paradigm. *Journal of Abnormal Child Psychology, 35*(5), 745–758.

Bálint, S., Czobor, P., Komlósi, S., Meszaros, A., Simon, V., & Bitter, I. (2009). Attention deficit hyperactivity disorder (ADHD): Gender-and age-related differences in neurocognition. *Psychological Medicine, 39*(8), 1337–1345.

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

Barkley, R. A. (2015). No Title. In R. A. Barkley (Ed.), *Attention-Deficit Hyperactivity Disorder, Fourth Edition: A Handbook for Diagnosis and Treatment* (pp. 356–390). Guilford Press.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological), 57*(1), 289–300.

Bledsoe, J. C., Semrud-Clikeman, M., & Pliszka, S. R. (2013). Anterior cingulate cortex and symptom severity in attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 122*(2), 558.

Brocki, K. C., Eninger, L., Thorell, L. B., & Bohlin, G. (2010). Interrelations between executive function and symptoms of hyperactivity/impulsivity and inattention in preschoolers: A two year longitudinal study. *Journal of Abnormal Child Psychology, 38*(2), 163–171.

Buzy, W. M., Medoff, D. R., & Schweitzer, J. B. (2009). Intra-individual variability among children with ADHD on a working memory task: An ex-Gaussian approach. *Child Neuropsychology, 15*(5), 441–459.

Calub, C. A., Rapport, M. D., Friedman, L. M., & Eckrich, S. J. (2019). IQ and academic achievement in children with ADHD: The differential effects of specific cognitive functions. *Journal of Psychopathology and Behavioral Assessment, 41*, 639–651.

Campez, M., Raiker, J. S., Sarver, D. E., Friedman, L. M., Orban, S. A., & Rapport, M. D. (2020). Working Memory Capacity and ADHD Symptoms in Boys: Examining the Heterogeneity of Working Memory Functioning Using Latent Profile Analysis. *Journal of Psychopathology and Behavioral Assessment, 42*, 450–463.

Castellanos, F. X., Sonuga-Barke, E. J., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences, 10*(3), 117–123.

Castellanos, F. X., Sonuga-Barke, E. J., Scheres, A., Di Martino, A., Hyde, C., & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological Psychiatry, 57*(11), 1416–1423.

Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience, 3*(8), 617–628.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*(2), 163–183.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786.

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society, 15*(3), 331–343.

DeRonda, A., Zhao, Y., Seymour, K. E., Mostofsky, S. H., & Rosch, K. S. (2021). Distinct Patterns of Impaired Cognitive Control Among Boys and Girls with ADHD Across Development. *Research on Child and Adolescent Psychopathology, 49*(7), 835–848.

Dirlikov, B., Rosch, K. S., Crocetti, D., Denckla, M. B., Mahone, E. M., & Mostofsky, S. H. (2015). Distinct frontal lobe morphology in girls and boys with ADHD. *Neuroimage: Clinical*, *7*, 222–229.

DuPaul, G. J., Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Academic and social functioning associated with attention-deficit/hyperactivity disorder: Latent class analyses of trajectories from kindergarten to fifth grade. *Journal of Abnormal Child Psychology, 44*(7), 1425–1438.

Eckrich, S. J., Rapport, M. D., Calub, C. A., & Friedman, L. M. (2019). Written expression in boys with ADHD: The mediating roles of working memory and oral expression. *Child Neuropsychology, 25*(6), 772–794.

Emch, M., von Bastian, C. C., & Koch, K. (2019). Neural correlates of verbal working memory: An fMRI meta-analysis. *Frontiers in Human Neuroscience, 13*, 180.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309.

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., Brinkman, W. B., Froehlich, T., Simon, J. O., & Altaye, M. (2011). Evidence for higher reaction time variability for children with ADHD on a range of cognitive tasks including reward and event rate manipulations. *Neuropsychology, 25*(4), 427.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26*(2), 471–479.

Fassbender, C., Zhang, H., Buzy, W. M., Cortes, C. R., Mizuiri, D., Beckett, L., & Schweitzer, J. B. (2009). A lack of default network suppression is linked to increased distractibility in ADHD. *Brain Research, 1273*, 114–128.

Feige, B., Biscaldi, M., Saville, C. W. N., Kluckert, C., Bender, S., Ebner-Priemer, U., Hennighausen, K., Rauh, R., Fleischhaker, C., & Klein, C. (2013). On the temporal characteristics of performance variability in attention deficit hyperactivity disorder (ADHD). *PloS One*, *8*(10), e69674.

Friedman, L. M., Rapport, M. D., Calub, C. A., & Eckrich, S. J. (2018a). ADHD and Core Foundational Learning: Working Memory's Contribution to Reading Comprehension and Applied Math Problem-Solving Abilities. *The ADHD Report, 26*(7), 1–7.

Friedman, L. M., Rapport, M. D., Raiker, J. S., Orban, S. A., & Eckrich, S. J. (2017). Reading comprehension in boys with ADHD: The mediating roles of working memory and orthographic conversion. *Journal of Abnormal Child Psychology, 45*(2), 273–287.

Friedman, L. M., Rapport, M. D., Orban, S. A., Eckrich, S. J., & Calub, C. A. (2018b). Applied problem solving in children with ADHD: The mediating roles of working memory and mathematical calculation. *Journal of Abnormal Child Psychology, 46*(3), 491–504.

Gadow, K. D., & Sprafkin, J. N. (2002). *Child symptom inventory 4: Screening and norms manual*. Checkmate Plus.

Gallo, E. F., & Posner, J. (2016). Moving towards causality in attention-deficit hyperactivity disorder: Overview of neural and genetic mechanisms. *The Lancet Psychiatry, 3*(6), 555–567.

Galloway-Long, H., & Huang-Pollock, C. (2018). Using inspection time and ex-Gaussian parameters of reaction time to predict executive functions in children with ADHD. *Intelligence, 69*, 186–194.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(8), 1036–1045.

Golay, P., Fagot, D., & Lecerf, T. (2013). Against coefficient of variation for estimation of intraindividual variability with accuracy measures. *Tutorials in Quantitative Methods for Psychology, 9*(1), 6–14.

Goos, L. M., Crosbie, J., Payne, S., & Schachar, R. (2009). Validation and extension of the endophenotype model in ADHD patterns of inheritance in a family study of inhibitory control. *American Journal of Psychiatry, 166*(6), 711–717.

Halperin, J. M., & Schulz, K. P. (2006). Revisiting the role of the prefrontal cortex in the pathophysiology of attention-deficit/hyperactivity disorder. *Psychological Bulletin, 132*(4), 560.

Helps, S. K., Broyd, S. J., Bitsakou, P., & Sonuga-Barke, E. J. S. (2011). Identifying a distinctive familial frequency band in reaction time fluctuations in ADHD. *Neuropsychology, 25*(6), 711.

Hollingshead, A. B. (1975). *Four factor index of social status*. Yale University.

Hornstra, R., Van der Oord, S., Staff, A. I., Hoekstra, P. J., Oosterlaan, J., Van der Veen-Mulders, L., Luman, M., & van den Hoofdakker, B. J. (2021). Which Techniques Work in Behavioral Parent Training for Children with ADHD? A Randomized Controlled Microtrial. *Journal of Clinical Child & Adolescent Psychology*, 1–16.

IBM Corp. (2019). IBM SPSS Statistics for Mac, Version 26.0. Armonk, NY: IBM Corp.

Irwin, L. N., Soto, E. F., Chan, E. S. M., Miller, C. E., Carrington-Forde, S., Groves, N. B., & Kofler, M. J. (2021). Activities of daily living and working memory in pediatric attention-deficit/hyperactivity disorder (ADHD). *Child Neuropsychology, 27*(4), 468–490.

Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36*(4), 567–589.

Karalunas, S. L. (2010). *Consistently inconsistent: understanding intra-individual variability in ADHD*. State College: Penn State University (Doctoral Dissertation).

Karatekin, C. (2004). A test of the integrity of the components of Baddeley's model of working memory in attention-deficit/hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry, 45*(5), 912–926.

Kasper, L. J., Alderson, R. M., & Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): A meta-analytic review. *Clinical Psychology Review, 32*(7), 605–617.

Kaufman, A. S. (1997). *Comprehensive form manual for the Kaufman Test of Educational Achievement-Normative Update*. American Guidance Service.

Kaufman, A. S. (2004). *Kaufman test of educational achievement, (KTEA-II)*. American Guidance Service.

Kaufman, A. S. (2014). *Kaufman Test of Educational Achievement-Third Edition (KTEA-3)*. Bloomington, MN: Pearson.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., & Ryan, N. (1996). *Kiddie-Sads-present and Lifetime version (K-SADS-PL)*. University of Pittsburgh, School of Medicine.

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry, 60*(10), 1088–1097.

Kofler, M. J., Irwin, L. N., Soto, E. F., Groves, N. B., Harmon, S. L., & Sarver, D. E. (2019). Executive functioning heterogeneity in pediatric ADHD. *Journal of Abnormal Child Psychology, 47*(2), 273–286.

Kofler, M. J., Raiker, J. S., Sarver, D. E., Wells, E. L., & Soto, E. F. (2016). Is hyperactivity ubiquitous in ADHD or dependent on environmental demands? Evidence from meta-analysis. *Clinical Psychology Review, 46*, 12–24.

Kofler, M. J., Rapport, M. D., & Alderson, R. M. (2008). Quantifying ADHD classroom inattentiveness, its moderators, and variability: A meta-analytic review. *Journal of Child Psychology and Psychiatry, 49*(1), 59–69.

Kofler, M. J., Rapport, M. D., Bolden, J., Sarver, D. E., & Raiker, J. S. (2010). ADHD and working memory: The impact of central executive deficits and exceeding storage/rehearsal capacity on observed inattentive behavior. *Journal of Abnormal Child Psychology, 38*(2), 149–161.

Kofler, M. J., Rapport, M. D., Bolden, J., Sarver, D. E., Raiker, J. S., & Alderson, R. M. (2011). Working memory deficits and social problems in children with ADHD. *Journal of Abnormal Child Psychology, 39*(6), 805–817.

Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S. A., Friedman, L. M., & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clinical Psychology Review, 33*(6), 795–811.

Kofler, M. J., Sarver, D. E., Spiegel, J. A., Day, T. N., Harmon, S. L., & Wells, E. L. (2017). Heterogeneity in ADHD: Neurocognitive predictors of peer, family, and academic functioning. *Child Neuropsychology, 23*(6), 733–759.

Kuntsi, J., Oosterlaan, J., & Stevenson, J. (2001). Psychological mechanisms in hyperactivity: I response inhibition deficit, working memory impairment, delay aversion, or something else? *Journal of Child Psychology and Psychiatry, 42*(2), 199–210.

Lange, K. W., Reichl, S., Lange, K. M., Tucha, L., & Tucha, O. (2010). The history of attention deficit hyperactivity disorder. *ADHD Attention Deficit and Hyperactivity Disorders, 2*(4), 241–255.

Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica,* 104, 167–190.

Lijffijt, M., Kenemans, J. L., Verbaten, M. N., & van Engeland, H. (2005). A meta-analytic review of stopping performance in attention-deficit/hyperactivity disorder: Deficient inhibitory motor control? *Journal of Abnormal Psychology, 114*(2), 216.

Matsumoto, K., & Tanaka, K. (2004). Conflict and cognitive control. *Science, 303*(5660), 969–970.

Martin, J. D., Tsukahara, J. S., Draheim, C., Shipstead, Z., Mashburn, C. A., Vogel, E. K., & Engle, R. W. (2021). The visual arrays task: Visual storage capacity or attention control? *Journal of Experimental Psychology: General*, *150*(12), 2525–2551.

Mestdagh, M., Pe, M., Pestman, W., Verdonck, S., Kuppens, P., & Tuerlinckx, F. (2018). Sidelining the mean: The relative variability index as a generic mean-corrected variability measure for bounded variables. *Psychological Methods, 23*(4), 690.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*(1), 40–48.

Mostert, J. C., Hoogman, M., Onnink, A. M. H., van Rooij, D., von Rhein, D., van Hulzen, K. J., ... & Franke, B. (2018). Similar subgroups based on cognitive performance parse heterogeneity in adults with ADHD and healthy controls. *Journal of Attention Disorders*, *22*(3), 281–292.

Nigg, J. T. (1999). The ADHD response-inhibition deficit as measured by the stop task: Replication with DSM–IV combined type, extension, and qualification. *Journal of Abnormal Child Psychology, 27*(5), 393–402.

Orban, S. A., Rapport, M. D., Friedman, L. M., Eckrich, S. J., & Kofler, M. J. (2018). Inattentive behavior in boys with ADHD during classroom instruction: The mediating role of working memory processes. *Journal of Abnormal Child Psychology, 46*(4), 713–727.

Porrino, L. J., Rapoport, J. L., Behar, D., Sceery, W., Ismond, D. R., & Bunney, W. E. (1983). A naturalistic assessment of the motor activity of hyperactive boys: I. Comparison with normal controls. *Archives of General Psychiatry*, *40*(6), 681–687.

Raiker, J. S., Friedman, L. M., Orban, S. A., Kofler, M. J., Sarver, D. E., & Rapport, M. D. (2019). Phonological Working Memory Deficits in ADHD Revisited: The Role of Lower Level Information-Processing Deficits in Impaired Working Memory Performance. *Journal of Attention Disorders, 23*(6), 570–583.

Raiker, J. S., Rapport, M. D., Kofler, M. J., & Sarver, D. E. (2012). Objectively-measured impulsivity and attention-deficit/hyperactivity disorder (ADHD): Testing competing predictions from the working memory and behavioral inhibition models of ADHD. *Journal of Abnormal Child Psychology, 40*(5), 699–713.

Rapport, M. D. (1994). Attention Deficit Hyperactivity Disorder. In M. Hersen & V. B. Van Hasselt (Eds.), *Advanced Abnormal Psychology* (pp. 189–206). Plenum Press.

Rapport, M. D., Alderson, R. M., Kofler, M. J., Sarver, D. E., Bolden, J., & Sims, V. (2008). Working memory deficits in boys with attention-deficit/hyperactivity disorder (ADHD): The contribution of central executive and subsystem processes. *Journal of Abnormal Child Psychology, 36*(6), 825–837.

Rapport, M. D., Bolden, J., Kofler, M. J., Sarver, D. E., Raiker, J. S., & Alderson, R. M. (2009). Hyperactivity in boys with attention-deficit/hyperactivity disorder (ADHD): A ubiquitous core symptom or manifestation of working memory deficits? *Journal of Abnormal Child Psychology, 37*(4), 521–534.

Rapport, M. D., Murphy, A., & Bailey, J. S. (1980). The effects of a response cost treatment tactic on hyperactive children. *Journal of School Psychology, 18*(2), 98–111.

Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review, 33*(8), 1237–1252.

Rescorla, L. A., & Achenbach, T. M. (2004). *The Achenbach System of Empirically Based Assessment (ASEBA) for Ages 18 to 90 Years.*

Roberts, W., Milich, R., & Barkley, R. A. (2015). *Attention-deficit hyperactivity disorder* (4th ed.). The Guilford Press.

Russell, V. A., Oades, R. D., Tannock, R., Killeen, P. R., Auerbach, J. G., Johansen, E. B., & Sagvolden, T. (2006). Response variability in attention-deficit/hyperactivity disorder: A neuronal and glial energetics hypothesis. *Behavioral and Brain Functions, 2*(1), 1–25.

Sarver, D. E., Rapport, M. D., Kofler, M. J., Raiker, J. S., & Friedman, L. M. (2015). Hyperactivity in attention-deficit/hyperactivity disorder (ADHD): Impairing deficit or compensatory behavior? *Journal of Abnormal Child Psychology, 43*(7), 1219–1232.

Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D., Clasen, L., Evans, A., Giedd, J., & Rapoport, J. L. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences, 104*(49), 19649–19654.

Solanto, M. V., Abikoff, H., Sonuga-Barke, E., Schachar, R., Logan, G. D., Wigal, T., Hechtman, L., Hinshaw, S., & Turkel, E. (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: A supplement to the NIMH multimodal treatment study of AD/HD. *Journal of Abnormal Child Psychology, 29*(3), 215–228.

Sonuga-Barke, E., Bitsakou, P., & Thompson, M. (2010). Beyond the dual pathway model: Evidence for the dissociation of timing, inhibitory, and delay-related impairments in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(4), 345–355.

Sonuga-Barke, E. J., & Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: A neurobiological hypothesis. *Neuroscience and Biobehavioral Reviews, 31*(7), 977–986.

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., Clevenger, W., Davies, M., Elliott, G. R., & Greenhill, L. L. (2001). Clinical relevance of the primary findings of the MTA: Success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(2), 168–179.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon.

Tamm, L., Narad, M. E., Antonini, T. N., O'Brien, K. M., Hawk, L. W., & Epstein, J. N. (2012). Reaction time variability in ADHD: A review. *Neurotherapeutics, 9*(3), 500–508.

Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of attention-deficit/ hyperactivity disorder: A systematic review and meta-analysis. *Pediatrics, 135*(4), e994–e1001.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience, 3*(4), 255–274.

Wåhlstedt, C., Thorell, L. B., & Bohlin, G. (2009). Heterogeneity in ADHD: Neuropsychological pathways, comorbidity and symptom domains. *Journal of Abnormal Child Psychology, 37*(4), 551–564.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children* (5th ed.). Pearson.

Wells, E. L., Kofler, M. J., Soto, E. F., Schaefer, H. S., & Sarver, D. E. (2018). Assessing working memory in children with ADHD: Minor administration and scoring changes may improve digit span backward's construct validity. *Research in Developmental Disabilities, 72*, 166–178.

Wolraich, M. L., Hagan, J. F., Allan, C., Chan, E., Davison, D., Earls, M., Evans, S. W., Flinn, S. K., Froehlich, T., & Frost, J. (2019). Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics*, *144*(4), e20192528.