J Abnorm Child Psychol
DOI 10.1007/s10802-017-0338-x

Inattentive Behavior in Boys with ADHD during Classroom Instruction: the Mediating Role of Working Memory Processes

Sarah A. Orban [1] · **Mark D. Rapport** [1] · **Lauren M. Friedman** [1] · **Samuel J. Eckrich** [1] · **Michael J. Kofler** [2]

Abstract Children with ADHD exhibit clinically impairing inattentive behavior during classroom instruction and in other cognitively demanding contexts. However, there have been surprisingly few attempts to validate anecdotal parent/teacher reports of intact sustained attention during 'preferred' activities such as watching movies. The current investigation addresses this omission, and provides an initial test of how ADHD-related working memory deficits contribute to inattentive behavior during classroom instruction. Boys ages 8–12 (M = 9.62, SD = 1.22) with ADHD (n = 32) and typically developing boys (TD; n = 30) completed a counterbalanced series of working memory tests and watched two videos on separate assessment days: an analogue math instructional video, and a non-instructional video selected to match the content and cognitive demands of parent/teacher-described 'preferred' activities. Objective, reliable observations of attentive behavior revealed no between-group differences during the non-instructional video (d = −0.02), and attentive behavior during the non-instructional video was unrelated to all working memory variables (r = −0.11 to 0.19, ns). In contrast, the ADHD group showed disproportionate attentive behavior decrements during analogue classroom instruction (d = −0.71). Bias-corrected, bootstrapped, serial mediation revealed that 59% of this between-group difference was attributable to ADHD-related impairments in central executive working memory, both directly (ER = 41%) and indirectly via its role in coordinating phonological short-term memory (ER = 15%). Between-group attentive behavior differences were no longer detectable after accounting for ADHD-related working memory impairments (d = −0.29, ns). Results confirm anecdotal reports of intact sustained attention during activities that place minimal demands on working memory, and indicate that ADHD children's inattention during analogue classroom instruction is related, in large part, to their underdeveloped working memory abilities.

Keywords ADHD · Attention-deficit/hyperactivity disorder · Classroom attention · Working memory · Classroom instruction

✉ Mark D. Rapport
mdrapport@gmail.com

[1] Department of Psychology, University of Central Florida, 4111 Pictor Lane, Psychology Bldg 99, Orlando, FL 32816, USA

[2] Department of Psychology, Florida State University, 1107 W. Call Street, Tallahassee, FL 32306, USA

Attention-deficit/hyperactivity disorder (ADHD) is an early-onset, heterogeneous neurodevelopmental disorder that affects an estimated 5–7% of children and adolescents worldwide (Polanczyk et al. 2007; Willcutt 2012). The primary symptoms of the disorder—chronic and developmentally excessive inattentiveness, gross motor activity, and impulsiveness—are associated with a wide range of functional impairments at home, while interacting with peers, and at school (cf. Barkley 2014; Hinshaw 2002; McQuade and Hoza 2008; Normand et al. 2013).

The classroom-related difficulties experienced by children with ADHD are well documented and particularly disconcerting due to their early onset, compounding course, and inverse relations with coveted academic performance and achievement outcomes. Classroom difficulties serve as an impetus for most clinical referrals (APA 2013; Pelham et al. 2005) and include a wide range of disadvantageous behaviors based on in vivo and analogue classroom studies. Relative to their classmates, children with ADHD complete fewer assignments correctly (DuPaul and Stoner 2014; Molina et al. 2009; Rapport et al. 1994), display higher rates of disruptive

Published online: 19 August 2017

behavior (Lauth et al. 2006), solicit more negative attention from teachers and peers (Abikoff et al. 2002; Skansgaard and Burns 1998), and exhibit higher rates of gross motor activity (DuPaul and Rapport 1993; Porrino et al. 1983; Vile Junod et al. 2006). Children with ADHD are also more than twice as inattentive as their non-ADHD classmates during teacher-directed classroom instructional activities (d = 1.19–1.40; Imeraj et al. 2013; Kofler et al. 2008). This inattentive behavior is usually attributed to underlying deficits in sustained attention rather than escape/avoidance behavior associated with excess gross motor activity (Abikoff et al. 2002; Dally 2006; del Mar Bernad et al. 2016; Rabiner et al. 2000; Spira and Fischel 2005; Vile Junod et al. 2006), and particularly troublesome given the multifaceted nature of classroom instruction and its importance to children's learning (Huitt et al. 2009; Slavin 2012).

Clarifying the mechanisms and processes responsible for inattentive behavior in ADHD is critical given its association with a host of adverse long-term outcomes (Shaw et al. 2012), particularly given the limited long-term benefits of extant evidence-based treatments (Molina et al. 2009; Riddle et al. 2013) and evidence that treatment-related improvements in these symptoms dissipate within minutes to hours of discontinuing psychosocial and pharmacological interventions (Chronis et al. 2004). To that end, a promising approach involves identifying the contexts in which these children are *not* less attentive than their unaffected peers as an initial step toward identifying differences between these contexts and those in which ADHD-related inattention are well documented (Kofler et al. 2008). In particular, children with ADHD appear to experience minimal difficulty remaining attentive while engaging in 'high interest' activities such as watching movies, playing video games, or drawing based on empirical (Kofler et al. 2010; Rapport et al. 2009) and parent/teacher anecdotal reports (Roberts et al. 2015).

Theoretical accounts of this phenomenon highlight the preferred vs. non-preferred nature of these activities, and posit that motivational deficits may underlie the decrement in attentive behavior observed during non-preferred activities (Luman et al. 2005; Sergeant, Oosterlaan, & van der Meere, 1999). Support for a motivational deficit is lacking, however, as most incentive studies fail to show disproportional improvement for children with ADHD relative to typically developing children, and even substantial rewards fail to normalize attentive behavior in children with ADHD (Dovis et al. 2012).

Alternative theoretical models (Rapport et al. 2008, 2009) and position papers (Lui and Tannock 2007; Martinussen and Tannock 2006) posit that the discrepancy between attentive behavior in children with ADHD during preferred and non-preferred activities may be better explained by differences in the neurocognitive processes required to successfully perform these activities (Kasper et al. 2012). For example, the ADHD working memory (WM) model (Rapport et al. 2008, 2009) posits that inattentive behavior during teacher-directed instruction reflects, in large part, an outcome of task demands that tax multiple, interacting WM processes—viz., the *working* and *memory* components of the WM system. The working component consists of a domain-general, frontally/prefrontally-mediated, central executive (CE) attentional controller that is responsible for updating and reordering internally-held information. It also provides oversight of two, modality-specific, anatomically distinct, short-term *memory* components and coordinates their interaction with information accessed from long-term memory (Baddeley 2007). The phonological short-term memory (PH STM) subsystem, localized in the left temporoparietal region and Broca's area, is responsible for the temporary storage and maintenance of verbal and written material that requires language-based processing, whereas the visuospatial short-term memory (VS STM) subsystem, localized in the dorsolateral/ventrolateral prefrontal cortex and posterior parietal/superior occipital cortices, provides this function for non-verbal and spatial information (cf. Baddeley 2003, 2007; Todd and Marois 2004). A third storage component—the episodic buffer—has also been proposed to handle multidimensional representations (Baddeley 2007); however, it was not examined in the current investigation due to our use of tasks that do not require binding multidimensional stimuli, coupled with recent evidence that it is not yet distinguishable from the VS STM subsystem in school-aged children (Gray et al. 2017).

The ADHD WM model hypothesizes that children with and without ADHD can attend equally well while engaged in activities that place minimal demands on WM, but will exhibit higher rates of inattention during activities that require considerable CE and PH/VS STM resources (e.g., during teacher-directed classroom instruction of core foundational subjects such as math). Under these latter conditions, children with ADHD are predicted to exhibit comparatively higher rates of inattentiveness due to their CE, and to a lesser extent, PH/VS STM deficits (Kasper et al. 2012; Kofler et al. 2010) consistent with an *in vivo* classroom observation study wherein children with ADHD were more inattentive while engaged in demanding academic tasks such as mathematics relative to music and art activities (Imeraj et al. 2013). These findings are expected given the well-established relations between WM and mathematical ability in children without ADHD (Swanson and Beebe-Frankenberger 2004; Swanson and Jerman 2006; Swanson and Kim 2007), as well as experimental evidence that increasing WM demands is associated with differential increases in inattentive behavior for children with ADHD relative to typically developing children (Kofler et al. 2010). Poorer performance on orally presented math problems that require CE-updating has also been demonstrated in children with teacher-rated ADHD symptoms (Re et al. 2016); however, its interplay with attention has not been elucidated. In addition, to our knowledge there have been no controlled studies examining whether ADHD-related inattentive behavior is magnified during math instruction relative to 'preferred' activities (e.g., watching a movie), or examining relations between WM components and attentive behavior during classroom instruction.

In summary, there have been surprisingly few attempts to empirically validate the anecdotal but oft-reported observation that attention deficits in ADHD are context dependent—that is, that these children demonstrate developmentally appropriate sustained attention during 'preferred' activities, but clinically impairing inattention during classroom instruction and other cognitively-demanding contexts. The current investigation addresses this omission, and provides an initial examination into the extent to which ADHD-related inattentive behavior during classroom instruction is related to their well-documented WM deficits (Kasper et al. 2012). A classroom analogue using an unconstrained natural viewing paradigm was used due to the impracticality of assessing both attention and WM component processes in individual children within an *in vivo* classroom environment. ADHD and typically developing (TD) boys were expected to exhibit similar, high rates of attentive behavior while watching a non-instructional, cognitively undemanding video. Both groups were expected to show significant decreases in attentive behavior while viewing a math instructional video, with disproportionate decreases for the ADHD relative to TD group.

After confirming that ADHD-related inattentive behavior is detectable only during the analogue classroom instruction video, a second set of analyses will test model-driven predictions that WM abilities would mediate these differences. PH STM was hypothesized to partially mediate the diagnostic status/task attention relation during the instructional video based on evidence that it plays a more limited role during math instruction (Friso-van den Bos et al. 2013; Swanson and Kim 2007). No hypothesis was proposed regarding the role of VS STM due to a lack of consensus in the field. CE processes associated with updating information and controlling interference (CE-updating) were hypothesized to fully mediate the diagnostic status (ADHD, TD) to task attention relation, based on previous evidence linking these abilities with children's skill at following instructions (Yang et al. 2014). In contrast, CE-reordering processes were not expected to explain incremental variance in the ADHD-attentive behavior relation given the lack of face-valid demands on this process while listening to math instructions. Finally, a serial mediation model was planned to test the hypothesis that the mediating role of CE-updating could be further parsed into unique and interactive effects with PH STM, based on the Baddeley (2007) conceptualization that the CE exerts oversight and coordination of the lower-level PH short-term storage subsystem. If detected, this finding would support WM model predictions that CE-mediated attentional focus enables successful updating and manipulation of PH STM contents while engaged in math instructional activities.

Method

Participants

The sample included 62 boys aged 8 to 12 years (M = 9.62, SD = 1.22) recruited by or referred to a children's learning clinic through community resources. Sample ethnicity was mixed and included 43 Caucasian non-Hispanic (69%), 13 Hispanic English-speaking (21%), 2 African American (3%), and 4 children of mixed racial/ethnic background (7%). All parents and children provided their informed consent/assent to participate in the study, and the university's Institutional Review Board approved the study prior to the onset of data collection. A psychoeducational evaluation was provided to the parents of all participants. Boys 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. A total of 21 children were excluded from the ADHD group over the 6-year period in which the study was conducted. Excluded children met diagnostic criteria for: intellectual disability (n = 2), unipolar depression (n = 5), generalized anxiety disorder (n = 4), autism spectrum disorder (n = 2), seizure disorder (n = 1), ADHD-Inattentive presentation (n = 4), ADHD-HI presentation (n = 1) and ADHD-NOS presentation (n = 1).

Group Assignment

All children and their parents participated in a detailed, semi-structured clinical interview using the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS; Kaufman et al. 1997). The K-SADS assesses onset, course, duration, severity, and impairment of current and past episodes of psychopathology in children based on DSM-IV criteria. Its psychometric properties are well established, including inter-rater agreement of 0.93 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. 1997).

Thirty-two children were included in the ADHD-Combined Type group based on: (1) an independent diagnosis by the directing clinical psychologist using DSM-IV criteria for ADHD-Combined Type[1] 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 and Rescorla 2001), or exceeding the criterion score for the parent version of the ADHD-Combined subtype subscale of the Child Symptom

[1] All children meeting DSM-IV criteria for ADHD-Combined Type met criteria using DSM-5 criteria for ADHD Combined Presentation.

Inventory-4: Parent Checklist (CSI-P; Gadow et al. 2004); and (3) teacher ratings of at least 2 *SD*s above the mean on the Attention-Deficit/Hyperactivity Problems DSM-Oriented scale of the CBCL Teacher Report Form (TRF), or exceeding the criterion score for the teacher version of the ADHD-Combined subtype subscale of the Child Symptom Inventory-4: Teacher Checklist. Fourteen (44%) of the children with ADHD were prescribed psychostimulants, which were withheld for 24-h prior to each testing session. Seven (22%) children also met criteria for Oppositional Defiant Disorder (ODD).

Thirty children were included in the TD group based on: (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) non-clinical range CSI subscale parent and teacher ratings. Typically developing children were recruited from a variety of resources including the clinic's webpage, information bulletins sent to neighborhood schools, and word of mouth.

Procedures

All tasks were administered as part of a larger battery that required the child's presence for approximately 2.5 h per session across four consecutive assessment sessions 1-week apart. Children completed all tasks while seated alone, approximately 0.66 m from a computer monitor. Performance was monitored at all times by an examiner stationed just outside the child's view to provide a structured setting while minimizing performance improvements associated with examiner demand characteristics (Power 1992). All children received brief (2–3 min) breaks following each task, and preset, longer (10–15 min) breaks after every two to three tasks.[2]

Working Memory

Phonological Working Memory (PH WM) The PH WM number-letter reordering task assesses PH WM based on Baddeley's (2007) model, and its cognitive demands require an active interplay between higher-order CE processes (attention and interference control, reordering) and subsidiary PH STM processes. Children were presented a series of jumbled numbers and a capital letter on a computer monitor. The letter never appeared in the first or last position to minimize primacy and recency effects, and was counterbalanced across trials to appear an equal number of times in the other serial positions. Children were instructed to recall the numbers in order from smallest to largest, and to say the letter last (e.g., 4 H 6 2 is correctly recalled as 2 4 6 H). Two trained research assistants, blind to diagnostic status and seated out of the child's view, recorded children's verbal responses independently on a pre-formatted response sheet. Inter-rater reliability was 96.3%; discrepancies were resolved via audio-video review. Previous studies have reported strong reliability and validity of the PH WM task, evidenced by high internal consistency ($r = 0.82$ to 0.97) and significantly large correlations ($r = 0.50$ to 0.71) with an established measure of working memory (i.e., WISC-IV Working Memory Index), respectively (Alderson et al. 2015; Raiker et al. 2012).

Visuospatial Working Memory (VS WM) The VS WM task is based on Baddeley's (2007) model, and its cognitive demands require an active interplay between upper level CE processes (i.e., attentional control and interference control, reordering) and subsidiary VS STM processes. Children were shown nine 3.2 cm squares arranged in three vertical columns on a computer monitor. The columns were offset from a standard 3 × 3 grid to minimize the likelihood of phonological coding of the stimuli (e.g., by equating the squares to numbers on a telephone pad). A series of 2.5 cm diameter dots were presented sequentially in one of nine squares during each trial, such that no two dots appeared in the same square on a given trial. All but one dot presented within the squares was black—the exception being a red dot that was counterbalanced across trials to appear an equal number of times in each of the nine squares, but never presented as the first or last stimulus to minimize primacy and recency. Children were instructed to respond by pressing the corresponding squares on a modified computer keyboard, and to re-order the dot locations by indicating the serial position of the black dots in the order presented followed by the serial position of the red dot last.

Five practice trials were administered before each PH and VS WM task (80% correct required). Each task involved 24 unique trials of the same set size, for eight total task conditions (set size 3–6, separately for PH and VS). Both tasks were independently counterbalanced across the four weekly assessment sessions, such that children received one PH and one VS task per session. Presentation rate was 800 ms per stimuli (200 ms inter-stimulus interval) for all PH and VS task variants. Evidence for reliability and validity of these working memory tasks includes high internal consistency ($\alpha = 0.82$), and demonstration of the expected magnitude of relations (Swanson and Kim 2007) with established measures of short-term memory (WISC-IV Digit Span raw scores: $r = 0.58$).

Working Memory/Serial Reordering Variables Partial-credit unit scoring (proportion of stimuli correct per trial)

[2] WM performance data for a subset of the current sample were used in separate studies to evaluate conceptually unrelated hypotheses (Alderson et al. 2010, 2012; Friedman et al. 2017a-online early release, b; Kofler et al. 2010, 2011, 2014; Raiker et al. 2012; Rapport et al. 2008, 2009; Sarver et al. 2012). We have not previously reported the instructional and non-instructional video data or their associations with our WM tasks for any children in the current sample.

was used as recommended (Conway et al. 2005). Estimates of central executive serial reordering (CE-reordering), phonological short-term memory (PH STM), and visuospatial short-term memory (VS STM) were computed at each set size using the procedures described by Rapport et al. (2008, 2009). Briefly, this involved regressing PH and VS performance at each set size onto each other to capture shared variance that reflects the domain-general, higher-order supervisory mechanism for the two processes. The final CE-reordering variable reflects a weighted average of these predicted scores based on their interrelations (i.e., factor score; CE-reordering factor loadings = 0.89–0.94), which has been shown to produce more accurate estimates of neurocognitive construct stability than confirmatory approaches (Willoughby et al. 2015). Similarly, the final PH STM (factor loadings = 0.56–0.74) and VS STM variables (factor loadings = 0.58–0.79) reflect the weighted average of their respective residual variances at each set size. Precedence for using shared variance to statistically derive CE-reordering and/or PH/VS STM variables is found for working memory components in Colom et al. (2005), Kane et al. (2004), Rosen and Engle (1997), and Swanson and Kim (2007).

Working Memory Updating The n-back task was designed to assess children's ability to temporarily store and continuously update information in working memory, and also requires controlled attention/interference control. The high-density, double-letter (1-back) n-back task described by Denney et al. (2005) and Raiker et al. (2012) was used in the current study (33.3% target density, 180 targets, 540 total stimuli, 200 ms presentation, 800 ms ISI). Children were instructed to press the mouse button every time a letter appeared that was the same as the previous letter (1-back), and to not respond to all other letter combinations. Total errors (both omission and commission errors) during the 9-min task served as the CE-updating independent variable. A practice block of 30 stimuli (10 targets) was administered (80% correct required). This task has been used previously by Raiker et al. (2012) and Denney et al. (2005) to examine whether WM updating ability may provide a more accurate conceptual explanation of children's performance errors that have traditionally been attributed to impulsivity and vigilance deficits. Recent studies of ADHD and TD children reveal large magnitude between-group differences on a similar n-back task (Alderson et al. 2017), and similar n-back tasks show moderate to high relations with complex WM span tasks ($r = 0.20$–0.97; Redick and Lindsey 2013; Schmiedek et al. 2014). Evidence for the task's reliability and validity includes high internal consistency ($r_{\text{block}} = 0.66$ to 0.90), expected relations with an AX (i.e., 'A' followed by 'X') version of the task (Denney et al. 2005), and high convergent validity with the CE-serial reordering variable used in the current investigation ($r = 0.73$).

Instructional and Non-instructional Video Clips

Video Clips Children were instructed to watch two, counterbalanced videos for 10-min each on separate assessment days. The video conditions were identical except for their content (e.g., same task instructions, audio volume, display size, testing room and chair). The *instructional video* was operationalized as an analogue to classroom instruction and featured a male instructor verbally and visually presenting multi-step solutions to addition, subtraction, and multiplication problems (e.g., notations when summing multiple addends that require a carry-over function). The video was selected for developmentally appropriate math content for our selected age range based on a standardized math skill assessment instrument (Good and Kaminski 2001). The *non-instructional video* featured the pod race scene from Star Wars Episode I, and was selected as an exemplar of the content and cognitive demands of 'preferred' activities during which children with ADHD reportedly demonstrate minimal attention deficits (i.e., rapidly changing scenes with no discernible manipulation/serial reordering and minimal short-term storage, rehearsal, or updating demands; Lui and Tannock 2007).

Direct Observations of Visual Attention Two trained observers, blind to children's diagnostic status, independently viewed and coded the video-recorded sessions using Observer XT 10.5 (Noldus Information Technology 2011). Observers completed extensive training and were required to obtain >80% agreement relative to a gold-standard prior to coding experimental data. Interrater reliability was assessed for all children across all tasks; percent agreement was 96.0%.

Visual attention was coded into one of two mutually exclusive states. Children were coded as *oriented* to task (i.e., attentive) when their head was directed within 45° vertically/horizontally of the center of the display screen. Children were coded as *not oriented* when their head direction exceeded a 45° vertical/horizontal tilt away from the screen's center. The *oriented* and *not oriented* codes are analogous to on- and off-task definitions used in most laboratory and classroom observation studies (Kofler et al. 2008). A continuous observation method with partial interval behavioral definitions was used to match previous ADHD classroom observation studies (Rapport et al. 2009). Behavioral states were changed (e.g., from oriented to not oriented) whenever the new behavioral state was present for ≥2 consecutive seconds.

Task attention was defined as the proportional duration children were visually oriented to the video screen during each of the two conditions (percent oriented). This frequency-based metric was selected to objectify children's attention while closely matching the frequency-based metric from most parent/teacher questionnaires. Support for the ecological validity of attentive behavior during these tasks includes significant associations with teacher-rated inattention for the instructional video ($r = -0.29$, $p = 0.03$) but not the non-instructional video ($r = -0.02$, $p = 0.86$), as well as recent results indicating similar rates of on-task behavior for ADHD and TD children during classroom math instruction (i.e., 81% and 91%, respectively; Imeraj et al. 2013).

Results

Data Screening

All independent and dependent variables were screened for multivariate (Mahalanobis distance $p < 0.001$) and univariate outliers (>3.0 *SD* from group mean). The PH STM factor score for one ADHD child was windsorized relative to the ADHD group mean as recommended (Tabachnick and Fidell 2007). TD group mean substitution was used for two TD children because they were homeschooled by the same informant who completed the parent forms (0.0002% of available data points); interpretation of results is unchanged if excluding these cases.

Data Analytic Overview

A three-tier analytic approach was adopted to examine the study's hypotheses. Preliminary analyses characterized the sample in terms of parent/teacher ratings, FSIQ, age, and SES. Tier 1 probed for the hypothesized group x condition interaction to investigate anecdotal reports that children with ADHD are less inattentive during classroom instruction but not 'preferred' activities. Tiers 2 and 3 used bias-corrected, bootstrapped mediation to examine the extent to which between-group differences in attentive behavior during math instruction were uniquely or jointly attributable to ADHD-related impairments in CE and PH/VS STM processes.

Preliminary Analyses

All parent and teacher ratings were higher for the ADHD relative to TD group as expected (Table 1). The groups did not differ in SES, $p = 0.12$. Children with ADHD, $M = 9.3$ years, $SD = 1.1$, were younger by about 2.2 months than TD children, $M = 9.9$, $SD = 1.3$; $p = 0.05$; age was therefore included as a covariate in all analyses. Between-group differences in FSIQ also reached significance, $p = 0.04$. FSIQ was not analyzed as a covariate, however, because it shares significant variance with working memory ($r = 0.75$ to 0.79; Weschler, 2007) and would result in removing substantial variance associated with WM from WM (Dennis et al. 2009). Consistent with past studies (e.g., Friedman et al. 2017b; Rapport et al. 2008, 2009), between-group differences in FSIQ were examined by removing reliable variance associated with the CE-serial reordering factor (described above) from FSIQ and then examining between-

Table 1 Sample and Demographic Variables

Variable	ADHD		Typically Developing		*F*	Cohen's *d*
	M	*SD*	*M*	*SD*		
Age	9.31	1.06	9.94	1.32	2.05*	−0.53
FSIQ	104.72	11.31	110.57	10.91	2.07*	−0.53
$FSIQ_{res}$	0.02	1.01	−0.03	1.00	−0.20	−0.05
SES	48.59	10.95	52.82	10.09	0.12	−0.40
CBCL AD/HD Problems	15.63	15.12	3.27	3.99	−4.34**	1.10
TRF AD/HD Problems	18.41	5.47	6.5	9.89	−5.92**	1.50
CSI-P: ADHD, Combined	38.28	9.05	9.93	9.69	−11.91**	3.02
CSI-T: ADHD, Combined	32.13	11.15	8.89	8.29	−9.26**	2.35
Phonological STM Factor Score	−0.26	1.12	0.28	0.79	2.18*	−0.55
Visuospatial STM Factor Score	−0.44	0.92	0.47	0.87	4.01**	−1.02
Central Executive Reordering Factor Score	−0.59	0.94	0.63	0.60	6.01**	−1.55
Central Executive Updating Z-score	−0.51	1.03	0.54	0.63	14.30**	−1.23

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

* $p \leq 0.05$, ** $p \leq 0.001$

group differences in FSIQ without the influence of the CE.[3] Initial group differences in FSIQ (Table 1) were no longer significant after residualizing for WM, $p = 0.93$ (see $FSIQ_{res}$ in Table 1).

Tier 1: Attentive Behavior by Group and Condition

Power Analysis G*Power v3.1 (Faul et al. 2007) a priori power analysis indicated that a total sample size of 55 is required to reliably detect between-group differences and interactions for power = 0.80, $\alpha = 0.05$, and two measurements (instructional and non-instructional videos) based on the expected $d = 1.40$ for observed classroom attentiveness (Kofler et al. 2008).

ANCOVA The 2 (TD, ADHD) × 2 (Math Instructional Video, Non-instructional Video) ANCOVA covaried for age, $p = 0.96$, revealed a significant main effect for group, $p = 0.01$, but not video condition, $p = 0.96$. There was a significant group x condition interaction, $p = 0.01$, that was attributable to disproportionate, cross-condition attentive behavior decreases for the ADHD group. Specifically, the ADHD, $M = 98.91$, $SD = 1.69$, and TD groups, $M = 98.94$, $SD = 1.69$, were highly attentive and not significantly different during the non-instructional video, $p = 0.94$; $d = -0.02$. In contrast, the ADHD group, $M = 83.66$, $SD = 12.86$ exhibited significant, large magnitude deficits in attention relative to TD controls, $M = 92.98$, $SD = 12.87$ during the math instructional video, $p = 0.01$; $d = -0.72$.

Tier 2: Simple Mediating Effects of WM Processes on ADHD/Attentive Behavior Relations

Power Bias-corrected, bootstrapped mediation requires a total $N = 34$ to reliably detect mediator effects of the expected magnitude for power = 0.80 and $\alpha = 0.05$ (Fritz and MacKinnon 2007), based on expected large associations between ADHD and WM (Fig. 1 path a; Kasper et al. 2012), WM and objectively observed attention (path b; Kofler et al. 2010), and ADHD and observed classroom attention (path c; Kofler et al. 2008). Thus, our $N = 62$ indicates adequate power.

Task Selection Mediation was not conducted for attention during the non-instructional video due to the lack of between-group differences and restricted range ($M = 99\%$ attentive for both groups). All WM components were impaired in ADHD (Table 1) and therefore retained as potential mediators of ADHD-related attentive behavior deficits during math instruction.

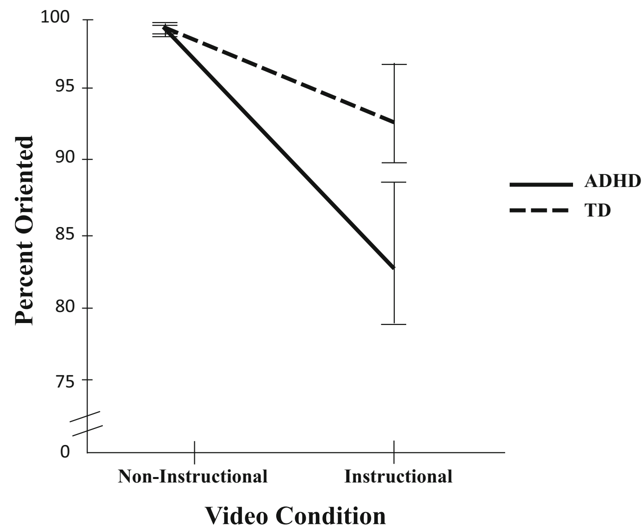

Fig. 1 Percent oriented for children with attention-deficit/hyperactivity disorder (ADHD; solid line) and typically developing children (TD; dashed line) during the Non-Instructional (Star Wars) and Instructional (Math) video conditions. *Vertical bars* represent standard deviation

Simple Mediation Overview Potential mediating effects of PH STM (Fig. 2a), VS STM (Fig. 2b), CE-reordering (Fig. 2c), and CE-updating (Fig. 2d) were tested initially, covaried for age. Continuous variables were converted to full-sample z-scores to allow unstandardized B weights to be interpreted as Cohen's d effect sizes when predicting from a dichotomous grouping variable (Hayes 2009). The PROCESS script for SPSS (Hayes 2014) was used for all analyses and 5000 samples were derived from the original sample ($N = 62$) by a process of resampling with replacement (Shrout and Bolger 2002). Ninety percent confidence intervals were selected to promote a more conservative evaluation of the extent to which inclusion of the mediating effect attenuates the direct effects of ADHD status on attentive behavior (Shrout and Bolger 2002).[4] Effect ratios (ER: indirect effect divided by total effect) were calculated to estimate the proportion of each significant total effect that was attributable to the mediating pathway. Direct effects are reported in Figs. 2, 3, and 4.

[3] Alternative approaches were considered but not adopted because they share substantial variance with WM (e.g., the WISC-IV General Ability Index (GAI) is comprised of the Verbal Comprehension and Perceptual Reasoning Indices, which shares 23% to 40% of variance with the WMI; Wechsler 2007).

[4] Briefly, the narrower 90% confidence interval is less likely to include 0.0, and represents a more conservative approach for estimating the magnitude of the relation between diagnostic status and the dependent variable after accounting for the mediator (i.e., partial mediation). For discussion and specific examples of this phenomenon, see Shrout and Bolger (2002). We acknowledge that 90% CIs are less conservative than 95% CIs for evaluating indirect pathways, despite being more conservative for the critical c' direct pathway. We emphasized the latter based on the rationale that Type II errors were most detrimental for the c' pathway (i.e., erroneously concluding that children with ADHD are not more inattentive than TD children after accounting for mediators). Current approaches preclude specifying different criteria for each pathway.

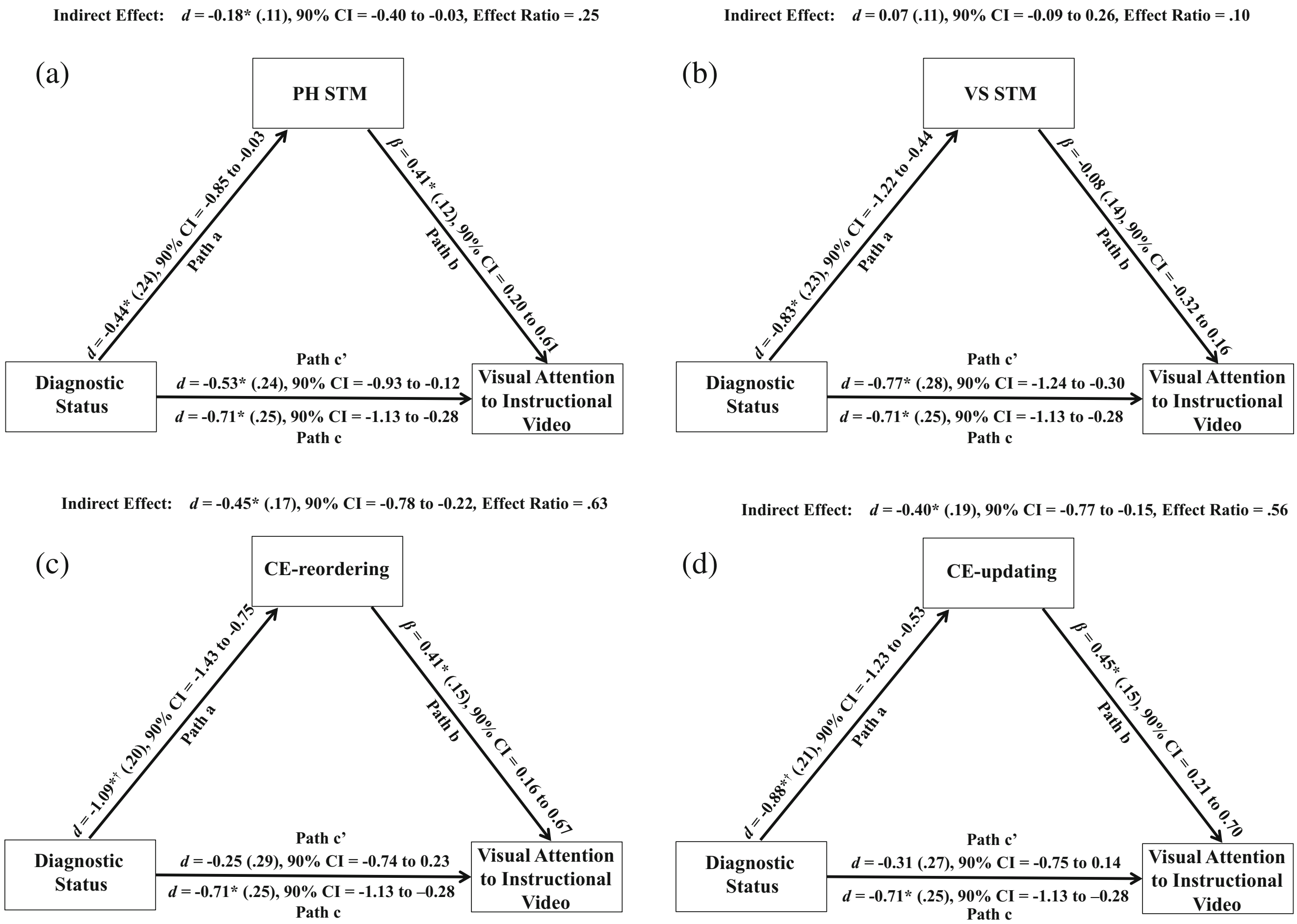

Fig. 2 Schematics depicting the effect sizes, standard errors and β coefficients of the total, direct, and indirect pathways for the mediating effect of (**a**) Phonological Short-Term Memory, (**b**) Visuospatial Short-Term Memory, (**c**) CE-reordering, and (**d**) CE-updating on Attentive behavior during the instructional video. Cohen's d for the c and c' pathways reflects the impact of ADHD diagnostic status on Attentive Behavior before (path c) and after (path c') taking into account the mediating variable. *Effect size (or β -weight) is significant based on 90% confidence intervals that do not include 0.0 (Shrout and Bolger 2002); values for path b reflect β -weights due to the use of two continuous variables in the calculation of the direct effect. † Path where age was a significant covariate

Indirect Effects of WM Components Indirect effects were significant for PH STM, $d = -0.18$, ER = 0.25, CE-reordering, $d = -0.45$, ER = 0.63, and CE-updating, $d = -0.40$, ER = 0.56, but not VS STM (90% CI includes 0.0). The effect ratios indicate that 25% (PH STM) to 56–63% (CE-reordering, CE-updating) of the ADHD-attentive behavior relation can be attributed to underdeveloped WM components in the ADHD group. These three WM components were retained for Tier 3, which considered these components together given their conceptual (Baddeley 2007) and statistical interrelations, $r = 0.62$–0.73.

Tier 3: Parallel and Serial Mediation of WM on ADHD/Attentive Behavior Relations

Tier 3 involved a 2-step process to determine the most parsimonious model for characterizing WM's association with ADHD-related inattention during the math instructional video. A parallel mediation model (Hayes 2014) that included both CE-reordering and CE-updating was used initially to examine whether their unique (reordering vs. updating) or shared (e.g., controlled attention, interference control) CE processes were responsible for their similar Tier 2 findings (Fig. 3). This model was predicated on their strong interrelations, $r = 0.73$, and meta-analytic evidence that they depend on both overlapping and non-overlapping prefrontal cortical structures (Nee et al. 2013; Wager and Smith 2003). We then tested for serial mediation (Hayes 2014), with the CE component from step 1 modeled to predict both attentive behavior and PH STM, and PH STM in turn also predicting attentive behavior (Fig. 4). This final model reflects the Baddeley (2007) conceptualization of the CE as responsible for reordering and updating information as well as oversight and coordination of the subsidiary PH STM subsystem (Fassbender and Schweitzer 2006; Luck et al. 2010).

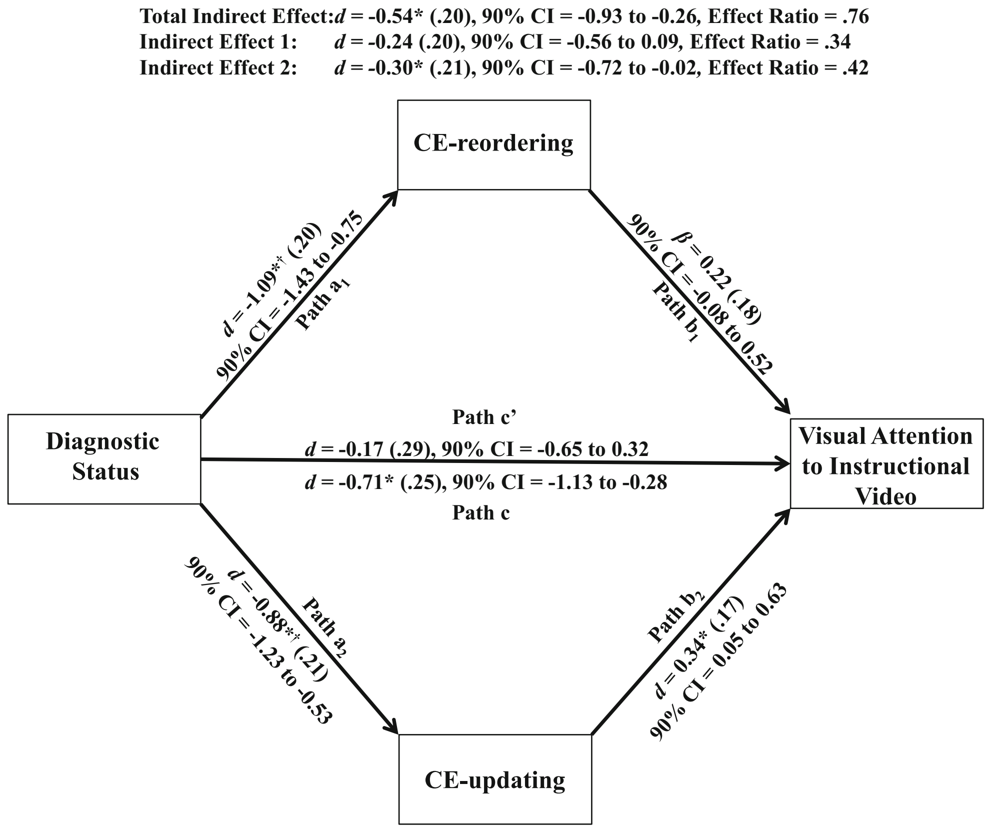

Fig. 3 Effect sizes, standard errors, and β coefficients of the total, direct, and indirect pathways for parallel mediation of CE-reordering and CE-updating on the relationship between Diagnostic Status and Attentive Behavior during the instructional video. Indirect Effect 1: Mediating effect of CE-reordering independent of Central Executive Updating on Attentive Behavior. Indirect Effect 2: Mediating effect of CE-updating independent of the CE-reordering on Attentive Behavior. Total Indirect Effect: Collective influence of both mediation pathways. CE: Central Executive. *Indicates significant path (90% CI does not include 0). † Path where age was a significant covariate

Parallel Mediation Inspection of Fig. 3 indicates that CE-updating, $d = -0.30$, ER = 0.42, but not CE-reordering, 90% CI includes 0.0, explained unique variance in the ADHD/attentive behavior relation.

Serial Mediation Based on the parallel mediation findings, CE-updating was tested using an exploratory serial mediation model with PH STM. Serial mediation allows the Tier 2 findings regarding CE-updating's mediating effect on the ADHD/attentive behavior relation to be further parsed into variance attributable to CE-updating specifically (Fig. 4, Indirect Effect 1) and CE-updating's role in governing the PH STM subsystem (Indirect Effect 3) while also considering potential unique PH STM influences (Indirect Effect 2).

In Tier 2, CE-updating's indirect effect was $d = -0.40$, and explained 56% of the ADHD/attentive behavior relation (Fig. 2d). As shown in Fig. 4, this effect can be further parsed into direct-indirect effects of CE-updating specifically, $d = -0.29$, ER = 0.41, and indirect-indirect effects of CE-updating via its role in governing PH STM, $d = -0.11$, ER = 0.15. Of note, these sub-indirect (serial mediation) effects will necessarily sum to the overall indirect effect reported in Tier 1 (i.e., $d = -0.29$ and -0.11 sum to $d = -0.40$, and ER = 0.41 + 0.15 = 0.56). Conceptually, the serial mediation model results provide preliminary evidence of the mechanisms by which the overall mediating effect operates and indicate that the ADHD group's large-magnitude deficits in attentive behavior during the math instructional video reflect, to a large extent, ADHD-related deficits in CE-updating abilities that facilitate engagement in complex instructional activities.

Discussion

The current study was the first to empirically demonstrate oft-reported yet anecdotal reports that children with ADHD 'can pay attention when they want to,' as evidenced by perceived TD-like sustained attention during 'preferred,' non-instructional activities and impaired attention during 'non-preferred,' academic instruction (Lui and Tannock 2007). An experimental, analogue methodology was adopted to permit more rigorous investigation of study hypotheses and involved objective, reliable observations of boys with and without ADHD while they watched two, counterbalanced videos selected to mirror *preferred*, high attention contexts and *non-preferred*, low attention academic instruction. Results revealed that both ADHD and TD children were highly attentive ($M = 99\%$ attentive) while viewing the non-instructional video, and significantly less attentive during the math instructional video. The hypothesized interaction effect was also supported: boys with ADHD demonstrated high rates of attention that did not differ from TD boys during the non-instructional video, but showed differential decreases during the math instructional video. Particularly noteworthy was the finding that the

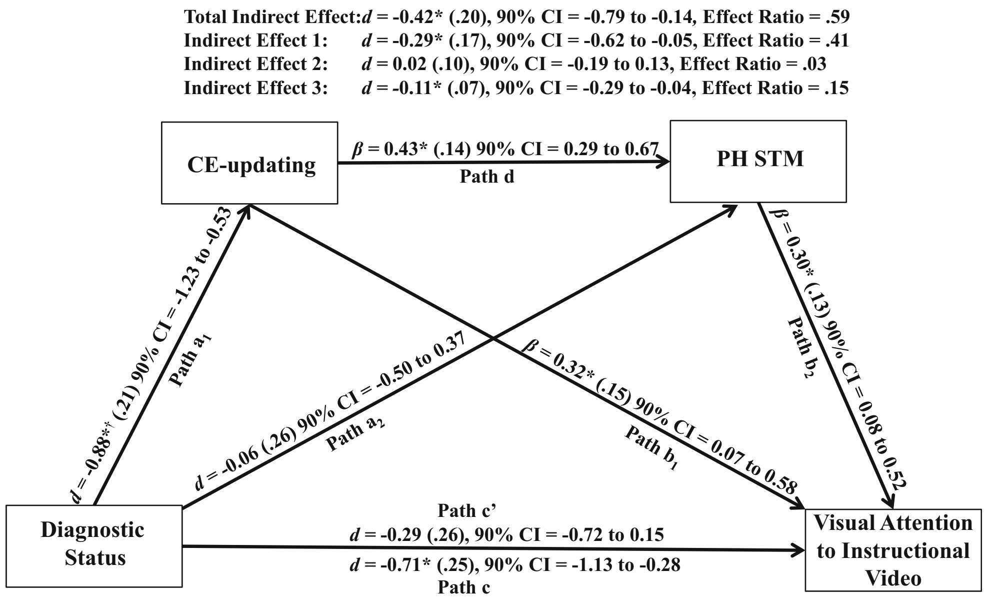

Fig. 4 Effect sizes, standard errors, and β coefficients of the total, direct, and indirect pathways for serial mediation of CE-updating and PH STM on the relationship between Diagnostic Status and Attentive Behavior during the instructional video. Indirect Effect 1: Mediating effect of CE-updating independent of PH STM on Attentive Behavior. Indirect Effect 2: Mediating effect of PH STM independent of CE-updating on Attentive Behavior. Indirect Effect 3: Mediating effect of the shared influence of CE-updating and PH STM on Attentive Behavior. Total Indirect Effect: Collective influence of all three mediation pathways. CE: Central Executive; PH STM: Phonological Short-Term Memory. *Indicates significant path (90% CI does not include 0). † Path where age was a significant covariate

attentive behavior during the non-instructional video was unrelated to all assessed WM processes.

Overall, our findings were consistent with past investigations in demonstrating that situational context (e.g., noise level, instructional delivery (Whalen et al. 1979), instructional communication cues (Zentall and Zentall 1983), and cognitive/executive function demands (Kofler et al. 2016, 2010; Rapport et al. 2009)) influence the display of core ADHD symptoms such as inattentiveness. It was the first study, however, to confirm anecdotal observations regarding intact sustained attention in ADHD during non-academic, cognitively undemanding activities, and demonstrate that maintaining high levels of attention varies according to video content and corresponding WM demands.

Of particular interest in the current study was the extent to which WM component processes were associated with attentive behavior during analogue classroom instruction, and the extent to which ADHD-related deficits in these WM components accounted for ADHD-related deficits in attentive behavior. Mediation analyses revealed that VS STM's contribution to attentive behavior was negligible and failed to mediate between-group attentive behavior differences. These findings were largely anticipated given minimal face-valid requirements to store visuospatial information during the math video. In contrast, CE-updating, CE-reordering, and PH STM emerged as significant mediators when modeled separately, but are more parsimoniously portrayed as interacting processes (Swanson and Fung 2016) based on the final, serial mediation model. Specifically, CE-reordering's effect was attributable to general CE processes rather than specific reordering demands, whereas CE-updating and PH STM act in tandem to fully attenuate between-group differences in attention during the instructional video. The finding that CE processes accounted for 56%–63% of ADHD-related inattentive behavior was striking, particularly given that CE abilities were assessed using three separate tasks that were distinct from the math video and administered on separate testing days.

Notably, children were not explicitly told to solve the math problems presented in the instructional video; however, verbally presented information gains automatic access to the PH STM subsystem, where it becomes immediately available for CE processing (Baddeley 2007). The strong link between CE abilities and attentive behavior during math instruction, combined with CE-updating's and CE-reordering's similar utility for explaining ADHD-related inattentive behavior, suggests that domain-general central executive functions are important for maintaining engagement when listening to and viewing teacher-directed, educational instruction. These CE functions include providing attentional control and oversight for the active process of updating needed information from long-term memory (e.g., math-related numbers, rules and algorithms) into the PH STM store, integrating this information with newly presented information, and removing unneeded information from PH STM to free-up space for additional information needed to keep track of the instructional content. The findings may also reflect, in part, underdeveloped CE-related interference control, which would allow irrelevant internal and/or external information to gain access to and interfere with the maintenance of instructional information in PH STM (Swanson and Fung 2016).

Deficiencies in the ability to update streaming information and process it continuously over a sustained duration—a prerequisite for comprehension in most educational tasks—appears particularly difficult for children with ADHD ($d = -0.88$) and results in losing critical information needed to pursue task goals. At these times, children are more likely to shift their attentional focus to irrelevant internal thoughts or external stimuli within the classroom (i.e., appear inattentive), consistent with the higher rates of attentional shifts (Rapport et al. 2009) and lower rates of attention (Kofler et al. 2008) observed for children with ADHD in classroom studies. Alternatively, basic attentional control may be limited in children with ADHD secondary to default mode network dysfunction (e.g., Fassbender et al. 2009), which intrudes on task-related thoughts while listening to teacher-directed instructions. This interpretation is consistent with our finding of a significant relation between PH STM and ADHD-related inattention. Meta-analytic evidence, however, has generally failed to support expected specificity of ADHD-related modulations at default mode frequencies (Karalunas et al. 2014; Kofler et al. 2013), and previous studies indicate that large-magnitude CE deficits in ADHD remain after accounting for their concurrently assessed attentive behavior (Kofler et al. 2010).

The involvement of the higher-order CE and subsidiary PH STM systems is consistent with past investigations of non-ADHD samples, but is the first to demonstrate this effect in children with ADHD and highlights the importance of CE updating processes for keeping track of classroom instructions. For example, Engle et al. (1991) found that PH WM (i.e., CE and PH STM measured as a single metric) and PH STM both predicted TD children's ability to follow oral instructions, with PH WM playing an increasingly important role as children progress from 1st to 6th grade. In contrast, two recent studies found that PH STM, rather than PH WM, showed the strongest continuity with children's success at following verbally-presented, multi-step instructions during in vivo (Gathercole et al. 2008) and virtual classrooms (Jaroslawska et al. 2016); however, neither study incorporated measures of CE-updating or examined whether CE and PH STM worked interactively. Similarly, Yang et al. (2014) tested TD children's memory for verbally presented instructions while engaged in a demanding secondary task intended to disrupt CE and PH STM processes. The resulting, large magnitude decrements in recall were consistent with the current findings, and indicate that both CE and PH STM processes are needed for children to update and maintain verbal instructions.

Limitations

Despite methodological (e.g., multiple tasks to estimate WM related PH/VS STM and CE) and statistical (e.g., bootstrapped mediation) refinements, limitations are inherent to all research investigations. Due to the well-documented gender differences related to ADHD primary symptom prevalence and course (Gaub and Carlson 1997; Williamson and Johnston 2015), neurocognitive functioning (Bálint et al. 2008), and neural morphology (Dirlikov et al. 2015), the current study focused exclusively on boys. Replication using larger, more diverse samples of children that include girls, adolescents, and additional ADHD presentation types is needed to examine the generalizability of the results. Additional benefit may also accrue by examining the extent to which the current findings extend to children diagnosed with other clinical disorders where WM and attentional deficits are suspected—e.g., autistic spectrum disorder (Luna et al. 2002; Swanson and Sachse-Lee 2001), internalizing disorders (Tannock et al. 1995) and externalizing disorders (Rhodes et al. 2012)— to elucidate the extent to which classroom inattentiveness during teacher-led instructional activities observed in children with other clinical disorders is due to similar mechanisms and processes as those reported for children with ADHD.

Children's attentive behavior during the math instructional video was marginally higher than rates reported for some in vivo classroom observational studies (Kofler et al. 2008), while other studies have found similar rates of attentive behavior for children with ADHD and TD children (Imeraj et al. 2013). Higher rates of attention in the current study may reflect the (a) absence of nearby children and customary distractions inherent to classroom settings; and/or (b) higher levels of expected frontal/prefrontal cortical activation and arousal associated with viewing and listening to movies documented via fMRI imaging (Vanderwal et al. 2017). Nevertheless, demonstrating the influence of WM processes on children's attention to instruction in a controlled experimental setting facilitated the dissection of the same underlying processes that likely operate in classroom settings.

Children's preexisting math knowledge may have influenced their attentive behavior during the math instructional video, particularly given the higher rates of math underachievement associated with ADHD (Frazier et al. 2007). Thus, although the video's instructional content was developmentally appropriate and the study was designed to minimize the influence of math skills (i.e., children were not instructed to perform any math calculations), the influence of children's behavioral learning histories cannot be ruled out. We considered controlling for math knowledge; however, this option was not feasible because approximately 70% of the variance in children's math test performance can be attributable to working memory processes (Swanson and Kim 2007). Additionally, the two videos differed in terms of rapidity in which scenes changed, the level of visual imagery and accompanying sounds, and other visual/auditory parameters. Future studies may benefit from varying the level of WM demands within exciting or highly stimulating video clips to determine if these factors play a role in ameliorating working memory deficits.

Finally, children with ADHD were off-task approximately 16% during the instructional video. While this estimate may appear inconsequential, the cumulative effects across weeks, months, and even years represents a significant loss of instructional time for children with ADHD compared to TD peers. Based on recommended education standards of total daily instructional time for mathematics among children in grades 3–5 (California Department of Education 2017), a 16% decrement in attention during the recommended 45-min daily math instructional period is equivalent to missing approximately 29 math lessons in an academic school year. This represents a potentially critical cumulative loss given the well-documented relationship between greater instructional time and coveted academic outcomes (Rivkin and Schiman 2015).

Clinical and Research Implications

The significant contributions of CE-updating and PH STM to children's attention during teacher-directed instructions have important implications for accommodating and remediating ADHD-related classroom behavior. For example, consideration of the congruence between an individual child's WM abilities and the WM demands of target classroom behaviors (e.g., maintaining attention during teacher-led instruction) may have important implications for determining reinforcement frequency and quantity. For example, children with greater WM deficits may require larger and/or more frequent rewards because the target behavior is objectively more difficult for them, although previous investigations indicate that even excessive incentives are inadequate (Dovis et al. 2012). More generally, compensatory interventions could involve re-structuring classroom activities to decrease the substantial WM demands associated with most instructional activities by incorporating mnemonics, cues, and visual aids to scaffold multi-step solutions, and separating multi-step instructions into independent steps. Compensatory classroom interventions for children with low WM, however, have been relatively unsuccessful to date (Colmar et al. 2016; Elliott et al. 2010) but may still hold promise. Similarly, working memory training programs are often touted as effective interventions for improving academic functioning in children with ADHD; however, multiple, independent meta-analytic reviews (Cortese et al. 2015; Melby-Lervåg and Hulme 2013; Rapport et al. 2013) uniformly indicate that these programs fail to promote clinically meaningful improvement in ecologically valid outcomes, including those related to classroom instruction and educational achievement.

Compliance with Ethical Standards

Funding This study was conducted without external funding.

Conflict of Interest The authors declare that they have no conflicts of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent/assent was obtained from all participants included in the study.

References

Abikoff, H. B., Jensen, P. S., Arnold, L. E., Hoza, B., Hechtman, L., Pollack, S., et al. (2002). Observed classroom behavior of children with ADHD: relationship to gender and comorbidity. *Journal of Abnormal Child Psychology, 30*(4), 349–359.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-age Forms & Profiles*. Burlington: University of Vermont.

Alderson, R. M., Rapport, M. D., Hudec, K. L., Sarver, D. E., & Kofler, M. J. (2010). Competing core processes in attention-deficit/hyperactivity disorder (ADHD): do working memory deficiencies underlie behavioral inhibition deficits? *Journal of Abnormal Child Psychology, 38*, 497–507.

Alderson, R. M., Rapport, M. D., Kasper, L. J., Sarver, D. E., & Kofler, M. J. (2012). Hyperactivity in boys with attention deficit/hyperactivity disorder (ADHD): the association between deficient behavioral inhibition, attentional processes, and objectively measured activity. *Child Neuropsychology, 18*, 487–505.

Alderson, R. M., Kasper, L. J., Patros, C. H., Hudec, K. L., Tarle, S. J., & Lea, S. E. (2015). Working memory deficits in boys with attention deficit/hyperactivity disorder (ADHD): an examination of orthographic coding and episodic buffer processes. *Child Neuropsychology, 21*(4), 509–530.

Alderson, R. M., Patros, C. H., 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.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Arlington: American Psychiatric Publishing.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Baddeley, A. (2007). *Working memory, thought, and action*. New York: Oxford University Press.

Bálint, S., Czobor, P., Komlosi, S., Meszaros, A., Simon, V., & Bitter, I. (2008). Attention deficit hyperactivity disorder (ADHD): gender- and age-related differences in neurocognition. *Psychological Medicine, 39*(8), 1337–1345.

Barkley, R. A. (Ed.). (2014). *Attention-deficit hyperactivity disorder: a handbook for diagnosis and treatment* (5th ed.). New York: Guilford Publications.

California Department of Education (2017). Recommended and required instructional minutes. Sacramento, CA. Retrieved from http://pubs.cde.ca.gov/tcsii/ch3/rdmdrqirdinstmin.aspx.

Chronis, A. M., Fabiano, G. A., Gnagy, E. M., Onyango, A. N., Pelham, W. E., Lopez-Williams, A., et al. (2004). An evaluation of the summer treatment program for children with attention-deficit/hyperactivity disorder using a treatment withdrawal design. *Behavior Therapy, 35*(3), 561–585.

Colmar, S., Davis, N., & Sheldon, L. (2016). A pilot classroom-based study of attention and working memory strategies for primary-aged students. *Journal of Psychologists & Counsellors in Schools, 26*, 125–136.

Colom, R., Abad, F. J., Rebollo, I., & Shih, P. C. (2005). Memory span and general intelligence: a latent-variable approach. *Intelligence, 33*(6), 623–642.

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.

Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., et al. (2015). Cognitive training for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*(3), 164–174.

Dally, K. (2006). The influence of phonological processing and inattentive behaviour on reading acquisition. *Journal of Educational Psychology, 98*(2), 420–437.

del Mar Bernad, M., Servera, M., Becker, S. P., & Burns, G. L. (2016). Sluggish cognitive tempo and ADHD inattention as predictors of externalizing, internalizing, and impairment domains: a 2-year longitudinal study. *Journal of Abnormal Child Psychology, 44*(4), 771–785.

Denney, C. B., Rapport, M. D., & Chung, K. M. (2005). Interactions of task and subject variables among continuous performance tests. *Journal of Child Psychology and Psychiatry, 46*(4), 420–435.

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.

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.

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. (2012). Can motivation normalize working memory and task persistence in children with attention-deficit/hyperactivity disorder? The effects of money and computer-gaming. *Journal of Abnormal Child Psychology, 40*(5), 669–681.

DuPaul, G. J., & Rapport, M. D. (1993). Does methylphenidate normalize the classroom performance of children with attention deficit disorder? *Journal of the American Academy of Child and Adolescent Psychiatry, 32*(1), 190–198.

DuPaul, G. J., & Stoner, G. (2014). *ADHD in the schools: assessment and intervention strategies*. New York: Guilford Publications.

Elliott, J. G., Gathercole, S. E., Alloway, T. P., Holmes, J., & Kirkwood, H. (2010). An evaluation of a classroom-based intervention to help overcome working memory difficulties and improve long-term academic achievement. *Journal of Cognitive Education and Psychology, 9*(3), 227–250.

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *The Journal of Educational Research, 84*(5), 253–262.

Fassbender, C., & Schweitzer, J. B. (2006). Is there evidence for neural compensation in attention deficit hyperactivity disorder? A review of the functional neuroimaging literature. *Clinical Psychology Review, 26*(4), 445–465.

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.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and achievement meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. *Journal of Learning Disabilities, 40*(1), 49–65.

Friedman, L.M., Rapport, M.D., Orban, S.A., Eckrich, S.E., Calub, C.A. (2017a-online early release). Applied problem solving in children with ADHD: the mediating roles of working memory and mathematical calculation. *Journal of Abnormal Child Psychology*. doi:10.1007/s10802-017-0312-7.

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

Friso-van den Bos, I., Van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: a meta-analysis. *Educational Research Review, 10*, 29–44.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18*(3), 233–239.

Gadow, K., Sprafkin, J., Salisbury, H., Schneider, J., & Loney, J. (2004). Further validity evidence for the teacher version of the child symptom inventory-4. *School Psychology Quarterly, 19*, 50–71.

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology, 22*(8), 1019–1037.

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

Good, R. H., & Kaminski, R. A. (Eds.). (2001). *Dynamic indicators of basic early literacy skills* (5th ed.). Eugene: Institute for the Development of Educational Achievement.

Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language, 92*, 183–201.

Hayes, A. F. (2009). Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408–420.

Hayes, A.F. (2014). PROCESS for SPSS (Version 2.12.1) (computer software). Columbus: The Guilford Press.

Hinshaw, S.P. (2002). Is ADHD an impairing condition in childhood and adolescence? In P.S. Jensen, & J.R. Cooper (Eds.), *Attention deficit hyperactivity disorder: state of the science—best practices* (pp. 5-1-5-21). Kingston: Civic Research Institute.

Huitt, W., Monetti, D., & Hummel, J. (2009). Designing direct instruction. In C. Reigeluth & A. Carr-Chellman (Eds.), *Instructional design theories and models: Volume III, building a common knowledgebase* (pp. 73–97). Mahwah: Lawrence Erlbaum Associates.

Imeraj, L., Antrop, I., Sonuga-Barke, E., Deboutte, D., Deschepper, E., Bal, S., & Roeyers, H. (2013). The impact of instructional context on classroom on-task behavior: a matched comparison of children with ADHD and non-ADHD classmates. *Journal of School Psychology, 51*(4), 487–498.

Jaroslawska, A. J., Gathercole, S. E., Allen, R. J., & Holmes, J. (2016). Following instructions from working memory: why does action at encoding and recall help? *Memory & Cognition, 44*(8), 1183–1191.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189–217.

Karalunas, S. L., Geurts, H. M., Konrad, K., Bender, S., & Nigg, J. T. (2014). Annual research review: reaction time variability in ADHD and autism spectrum disorders: measurement and mechanisms of a proposed trans-diagnostic phenotype. *Journal of Child Psychology and Psychiatry, 55*(6), 685–710.

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, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(7), 980–988.

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*, 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*, 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., Alderson, R. M., Raiker, J. S., Bolden, J., Sarver, D. E., & Rapport, M. D. (2014). Working memory and intraindividual variability as neurocognitive indicators in ADHD: examining competing model predictions. *Neuropsychology, 28*, 459–471.

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.

Lauth, G. W., Heubeck, B. G., & Mackowiak, K. (2006). Observation of children with attention-deficit hyperactivity (ADHD) problems in three natural classroom contexts. *British Journal of Educational Psychology, 76*(2), 385–404.

Luck, D., Danion, J. M., Marrer, C., Pham, B. T., Gounot, D., & Foucher, J. (2010). The right parahippocampal gyrus contributes to the formation and maintenance of bound information in working memory. *Brain and Cognition, 72*(2), 255–263.

Lui, M., & Tannock, R. (2007). Working memory and inattentive behaviour in a community sample of children. *Behavioral and Brain Functions, 3*(12), 1–11.

Luman, M., Oosterlaan, J., & Sergeant, J. A. (2005). The impact of reinforcement contingencies on AD/HD: a review and theoretical appraisal. *Clinical Psychology Review, 25*(2), 183–213.

Luna, B., Minshew, N. J., Garver, K. E., Lazar, N. A., Thulborn, K. R., Eddy, W. F., & Sweeney, J. A. (2002). Neocortical system abnormalities in autism: an fMRI study of spatial working memory. *Neurology, 59*(6), 834–840.

Martinussen, R., & Tannock, R. (2006). Working memory impairments in children with attention-deficit hyperactivity disorder with and without comorbid language learning disorders. *Journal of Clinical and Experimental Neuropsychology, 28*(7), 1073–1094.

McQuade, J. D., & Hoza, B. (2008). Peer problems in attention deficit hyperactivity disorder: current status and future directions. *Developmental Disabilities Research Reviews, 14*(4), 320–324.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270.

Molina, B.S., Hinshaw, S.P., Swanson, J.M., Arnold, L.E., Vitiello, B., Jensen, P.S., … MTA Cooperative Group. (2009). The MTA at 8 years: prospective follow-up of children treated for combined-type ADHD in a multisite study. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(5), 484–500.

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex, 23*(2), 264–282.

Noldus Information Technology (2011). *The observer XT reference manual version 10.5*. Wageningen: Author.

Normand, S., Schneider, B. H., Lee, M. D., Maisonneuve, M. F., Chupetlovska-Anastasova, A., Kuehn, S. M., & Robaey, P. (2013). Continuities and changes in the friendships of children with and without ADHD: a longitudinal, observational study. *Journal of Abnormal Child Psychology, 41*(7), 1161–1175.

Pelham Jr., W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 449–476.

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *American Journal of Psychiatry, 164*(6), 942–948.

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.

Power, T. J. (1992). Contextual factors in vigilance testing of children with ADHD. *Journal of Abnormal Child Psychology, 20*(6), 579–593.

Rabiner, D., Coie, J. D., & Conduct Problems Prevention Research Group. (2000). Early attention problems and children's reading achievement: a longitudinal investigation. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(7), 859–867.

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., Denney, C., DuPaul, G. J., & Gardner, M. J. (1994). Attention deficit disorder and methylphenidate: normalization rates, clinical effectiveness, and response prediction in 76 children. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*(6), 882–893.

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 attentiondeficit/hyperactivity disorder (ADHD): A ubiquitous core symptoms or manifestation of working memory deficits. *Journal of Abnormal Child Psychology, 37*, 521–534.

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.

Re, A. M., Lovero, F., Cornoldi, C., & Passolunghi, M. C. (2016). Difficulties of children with ADHD symptoms in solving

mathematical problems when information must be updated. *Research in Developmental Disabilities, 59*, 186–193.
Redick, T. S., & Lindsey, D. R. (2013). Complex span and n-back measures of working memory: a meta-analysis. *Psychonomic Bulletin & Review, 20*(6), 1102–1113.
Rhodes, S. M., Park, J., Seth, S., & Coghill, D. R. (2012). A comprehensive investigation of memory impairment in attention deficit hyperactivity disorder and oppositional defiant disorder. *Journal of Child Psychology and Psychiatry, 53*(2), 128–137.
Riddle, M. A., Yershova, K., Lazzaretto, D., Paykina, N., Yenokyan, G., Greenhill, L., et al. (2013). The preschool attention-deficit/hyperactivity disorder treatment study (PATS) 6-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(3), 264–278.
Rivkin, S. G., & Schiman, J. C. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal, 125*(588), F425–F448.
Roberts, W., Milich, R., & Barkley, R. (2015). Primary symptoms, diagnostic criteria, subtyping, and prevalence of ADHD. In R. Barkley (Ed.), *Attention-deficit hyperactivity disorder* (4th ed., pp. 50–80). New York: Guildford.
Rosen, V. M., & Engle, R. W. (1997). Forward and backward serial recall. *Intelligence, 25*, 37–47.
Sarver, D. E., Rapport, M. D., Kofler, M. J., Scanlan, S. W., Raiker, J. S., Altro, T. A., & Bolden, J. (2012). Attention problems, phonological short-term memory, and visuospatial short-term memory: differential effects on near-and long-term scholastic achievement. *Learning and Individual Differences, 22*, 8–19.
Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology, 5*, 1475.
Sergeant, J. A., Oosterlaan, J., & Van der Meere, J. (1999). Informationprocessing and energetic factors in attention-deficit/hyperactivity disorder. InH. C. Quay & A. E. Hogan (Eds.), Handbook of disruptive behavior disorders (pp. 75–104). New York: Kluwer Academic/Plenum Publishers.
Shaw, M., Hodgkins, P., Caci, H., Young, S., Kahle, J., Woods, A. G., & Arnold, L. E. (2012). A systematic review and analysis of long-term outcomes in attention deficit hyperactivity disorder: effects of treatment and non-treatment. *BMC Medicine, 10*(99), 1–15.
Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological Methods, 7*(4), 422–445.
Skansgaard, E. P., & Burns, G. L. (1998). Comparison of DSM-IV ADHD combined and predominantly inattention types: correspondence between teacher ratings and direct observations of inattentive, hyperactivity/impulsivity, slow cognitive tempo, oppositional defiant, and overt conduct disorder symptoms. *Child and Family Behavior Therapy, 20*, 1–14.
Slavin, R. E. (2012). *Educational Psychology: theory and practice: Edition 10*. Boston: Allyn & Bacon.
Spira, E. G., & Fischel, J. E. (2005). The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: a review. *Journal of Child Psychology and Psychiatry, 46*(7), 755–773.
Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*(3), 471.
Swanson, H. L., & Fung, W. (2016). Working memory components and problem-solving accuracy: are there multiple pathways? *Journal of Educational Psychology, 108*(8), 1153–1177.
Swanson, H. L., & Jerman, O. (2006). Math disabilities: a selective meta-analysis of the literature. *Review of Educational Research, 76*(2), 249–274.
Swanson, L., & Kim, K. (2007). Working memory, short-term memory, and naming speed as predictors of children's mathematical performance. *Intelligence, 35*(2), 151–168.
Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: both executive and phonological processes are important. *Journal of Experimental Child Psychology, 79*(3), 294–321.
Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon/Pearson Education.
Tannock, R., Ickowicz, A., & Schachar, R. (1995). Differential effects of methylphenidate on working memory in ADHD children with and without comorbid anxiety. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*(7), 886–896.
Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751–754.
Vanderwal, T., Eilbott, J., Finn, E.S., Craddock, R.C., Turnbull, A.G., Castellanos, F.X. (2017-online early release). Individual differences in functional connectivity during naturalistic viewing conditions. *NeuroImage* doi:10.1101/084665.
Vile Junod, R. E., DuPaul, G. J., Jitendra, A. K., Volpe, R. J., & Cleary, K. S. (2006). Classroom observations of students with and without ADHD: differences across types of engagement. *Journal of School Psychology, 44*(2), 87–104.
Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience, 3*(4), 255–274.
Wechsler, D. (2007). *Wechsler Intelligence Scale for Children—Fifth Edition*. Technical and interpretive manual. San Antonio: The Psychological Corporation.
Whalen, C. K., Henker, B., Collins, B. E., Finck, D., & Dotemoto, S. (1979). A social ecology of hyperactive boys: medication effects in structured classrooms. *Journal of Applied Behavior Analysis, 12*, 65–81.
Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics, 9*(3), 490–499.
Williamson, D., & Johnston, C. (2015). Gender differences in adults with attention-deficit/hyperactivity disorder: a narrative review. *Clinical Psychology Review, 40*, 15–27.
Willoughby, M. T., Blanton, Z. E., & Family Life Project Investigators. (2015). Replication and external validation of a bi-factor parameterization of attention deficit/hyperactivity symptomatology. *Journal of Clinical Child & Adolescent Psychology, 44*, 68–79.
Yang, T., Gathercole, S. E., & Allen, R. J. (2014). Benefit of enactment over oral repetition of verbal instruction does not require additional working memory during encoding. *Psychonomic Bulletin & Review, 21*, 186–192.
Zentall, S. S., & Zentall, T. (1983). Optimal stimulation: A model of disordered activity and performance in normal and deviant children. *Psychological Bulletin, 94,* 446–471.