



# Executive Functioning Heterogeneity in Pediatric ADHD

Michael J. Kofler<sup>1</sup> · Lauren N. Irwin<sup>1</sup> · Elia F. Soto<sup>1</sup> · Nicole B. Groves<sup>1</sup> · Sherelle L. Harmon<sup>1</sup> · Dustin E. Sarver<sup>2</sup>

© Springer Science+Business Media, LLC, part of Springer Nature 2018, corrected publication May/2018

## Abstract

Neurocognitive heterogeneity is increasingly recognized as a valid phenomenon in ADHD, with most estimates suggesting that executive dysfunction is present in only about 33%–50% of these children. However, recent critiques question the veracity of these estimates because our understanding of executive functioning in ADHD is based, in large part, on data from single tasks developed to detect gross neurological impairment rather than the specific executive processes hypothesized to underlie the ADHD phenotype. The current study is the first to comprehensively assess heterogeneity in all three primary executive functions in ADHD using a criterion battery that includes multiple tests per construct (working memory, inhibitory control, set shifting). Children ages 8–13 ( $M = 10.37$ ,  $SD = 1.39$ ) with and without ADHD ( $N = 136$ ; 64 girls; 62% Caucasian/Non-Hispanic) completed a counterbalanced series of executive function tests. Accounting for task unreliability, results indicated significantly improved sensitivity and specificity relative to prior estimates, with 89% of children with ADHD demonstrating objectively-defined impairment on at least one executive function (62% impaired working memory, 27% impaired inhibitory control, 38% impaired set shifting; 54% impaired on one executive function, 35% impaired on two or all three executive functions). Children with working memory deficits showed higher parent- and teacher-reported ADHD inattentive and hyperactive/impulsive symptoms ( $BF_{10} = 5.23 \times 10^4$ ), and were slightly younger ( $BF_{10} = 11.35$ ) than children without working memory deficits. Children with vs. without set shifting or inhibitory control deficits did not differ on ADHD symptoms, age, gender, IQ, SES, or medication status. Taken together, these findings confirm that ADHD is characterized by neurocognitive heterogeneity, while suggesting that contemporary, cognitively-informed criteria may provide improved precision for identifying a smaller number of neuropsychologically-impaired subtypes than previously described.

**Keywords** ADHD · Executive function · Working memory · Inhibition · Shifting · Heterogeneity

Attention-deficit/hyperactivity disorder (ADHD) is a chronic and impairing neurodevelopmental disorder that affects approximately 5% of school age-children (Polanczyk et al. 2014). Heterogeneity in the disorder's behavioral symptoms, associated impairments, and cognitive sequelae has been

increasingly recognized (e.g., Castellanos and Tannock 2002; Coghill et al. 2014; Sonuga-Barke et al. 2008). Current estimates suggest that only approximately 33%–50% of children with ADHD exhibit impairments in executive functioning (Biederman et al. 2004; Nigg et al. 2005). At the same time, recent critiques of the clinical literature's executive function task selection (e.g., Snyder et al. 2015) question the veracity of these estimates, such that our understanding of executive functioning in ADHD is based, in large part, on data from single tasks that may suboptimally assess their intended construct (Kofler et al. 2016; Coghill et al. 2014; Sonuga-Barke et al. 2008). The neurocognitive heterogeneity observed in ADHD is increasingly recognized as a valid phenomenon (Nigg et al. 2005); however, it is possible that cognitively-informed test batteries with greater construct precision and estimation with multiple tests may allow for the identification of a relatively small number of causal pathways to the ADHD

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10802-018-0438-2>) contains supplementary material, which is available to authorized users.

✉ Michael J. Kofler  
kofler@psy.fsu.edu

<sup>1</sup> Department of Psychology, Florida State University, 1107 W. Call Street, Tallahassee, FL 32306-4301, USA

<sup>2</sup> Departments of Pediatrics and Psychiatry, Center for Advancement of Youth, University of Mississippi Medical Center, Jackson, MS, USA

phenotype (Coghill et al. 2005). While previous investigations have assessed multiple neurocognitive constructs with single tasks, or included estimates of one or two of the three primary executive functions (for review, see Coghill et al. 2014), the current study is the first to comprehensively assess heterogeneity in all three primary executive functions in ADHD using a criterion battery that includes multiple tests per construct (working memory, inhibitory control, set shifting; Miyake et al. 2000).

## Executive Functioning

Executive functions refer to a set of interrelated, higher-order cognitive processes that enable goal directed behavior and novel problem solving (Baddeley 2007; Miyake et al. 2000). Among the diverse models of executive functioning, factor analytic and theoretical work provides the most empirical support for models that include three primary executive function domains: working memory, inhibitory control, and set shifting (Miyake et al. 2000; Miyake and Friedman 2012; St. Clair-Thompson & Gathercole 2006; Van der Sluis et al. 2007). These primary executive functions, in turn, enable goal-oriented behavior and support a host of secondary, non-executive cognitive abilities including but not limited to proactive and reactive interference control (Wiemers & Redick 2018) and goal-maintenance (Engle & Kane 2004), as well as performance on tasks intended to assess vigilance (Raiker et al. 2012), response variability (Kofler et al. 2014; Wiemers & Redick 2018), planning (Jaroslawska et al. 2016; Kofler et al. 2018; Miyake et al. 2000), perseveration (Miyake et al. 2000), and delay tolerance (Patros et al. 2015).

*Working memory* refers to the active, top-down manipulation of information held in short-term memory (Baddeley 2007), and includes interrelated functions of the mid-lateral prefrontal cortex and interconnected networks that involve updating, dual-processing, and temporal/serial reordering (Wager and Smith 2003). *Inhibitory control* refers to a set of interrelated cognitive processes that underlie the ability to withhold (action restraint) or stop (action cancellation) an on-going response (Alderson et al. 2007) and are supported by networks involving bilateral frontal, right superior temporal and left inferior occipital gyri, right thalamic, and mid-brain structures (Cortese et al. 2012). *Set shifting* refers to the ability to flexibly switch back-and-forth between mental sets via activation of prefrontal and posterior parietal cortices (Miyake et al. 2000; Pa et al. 2010).

## Executive Functioning in ADHD

Working memory and inhibitory control have been studied extensively in pediatric ADHD, with meta-analyses indicating

medium to large group-level impairments on tasks intended to assess these executive functions (Alderson et al. 2007; Kasper et al. 2012; Lijffijt et al. 2005; Martinussen et al. 2005; Willcutt et al. 2005). At the same time, a growing number of studies suggest that these group-level differences may be carried by impairments in a relatively small proportion of children with ADHD (Sonuga-Barke et al. 2008). Specifically, estimates for the proportion of pediatric ADHD cases who exhibit any form of executive dysfunction range from 21% to 60% across studies employing a wide range of tasks and impairment criteria (Biederman et al. 2004; Coghill et al. 2014; Fair et al. 2012; Geurts et al. 2006; Nigg et al. 2005; Solanto et al. 2001; Sonuga-Barke et al. 2010). Studies separating working memory and inhibitory control show similar heterogeneity: Working memory impairments are reported in 30% to 37% (Coghill et al. 2014; Fair et al. 2012), and inhibitory control deficits are detected in 21%–46% of pediatric ADHD cases (Coghill et al. 2014; Nigg et al. 2005; Solanto et al. 2001; Sonuga-Barke et al. 2010). To our knowledge, no ADHD study to date has examined heterogeneity in set shifting. However, meta-analytic effect sizes of  $d = 0.46$  to  $0.55$  (Willcutt et al. 2005) predict that 30%–36% of children with ADHD may demonstrate set shifting impairments based on converting effect size differences into expected population overlap proportions (Zakzanis 2001).

## Construct Validity and Task Impurity

These impairment estimates, combined with evidence that only a minority of ADHD cases are classified as impaired on multiple executive function tests (Biederman et al. 2004), have led in part to refined models of ADHD that emphasize causal heterogeneity (e.g., Nigg et al. 2005) and de-emphasize single-cause models that conceptualize ADHD as primarily a disorder of executive dysfunction (Coghill et al. 2005). However, as noted by Sonuga-Barke et al. (2008), these conclusions may be premature because the evidence-base is comprised, in large part, on data from clinical tests developed to detect gross neurological impairment rather than the more subtle deficits in cognitive control hypothesized to underlie the ADHD phenotype. That is, these tests may be most appropriate for screening for severe executive deficits in patients, but appear to lack sensitivity for targeting specific aspects of executive function and identifying individual differences across a wider range of abilities necessary for probing the nature of more subtle deficits associated with psychopathology (please see Snyder et al. 2015 for a review of specific tests).

**Construct validity and task specificity** The broad scope of the measurement issue raised by Sonuga-Barke et al. (2008) is highlighted when juxtaposing the tasks included in recent meta-analyses of ADHD neuropsychological functioning

(Alderson et al. 2007; Kasper et al. 2012; Martinussen et al. 2005; Willcutt et al. 2005) with a recent critique of executive function test selection in the clinical versus cognitive science literatures (Snyder et al. 2015). Across ADHD meta-analyses, estimates of deficits in working memory, inhibitory control, and set shifting are based in large part, and in some cases entirely, on data from tasks that have been criticized for sub-optimal construct validity (Redick and Lindsey 2013) and structural organization (Friedman and Miyake 2017) relative to contemporary advancements in cognitive psychology (Snyder et al. 2015). This suboptimal construct validity has demonstrated implications for our ability to identify executive dysfunction in children with ADHD. For example, Kasper et al. (2012) examined the evidence for working memory deficits in ADHD, and reported overall medium magnitude impairments ( $d = 0.69\text{--}0.74$ ) that were consistent with prior meta-analyses (Martinussen et al. 2005; Willcutt et al. 2005). At the same time, they found that the majority of tasks used to measure ‘working memory’ in the ADHD literature were better conceptualized as tests of ‘short term memory,’ with evidence from the cognitive literature suggesting that these tasks place relatively minimal demands on the executive components of working memory (Conway et al. 2005; Engle et al. 1999; Rapport et al. 2013). Using meta-regression, they found that their estimate of working memory deficits in ADHD increased from medium to very large ( $d = 2.0\text{--}2.2$ ) when based on working memory tasks with a prominent executive component congruent with contemporary cognitive definitions. Thus, whereas the uncorrected estimates suggest that 40%–44% of children with ADHD have working memory deficits (Zakzanis 2001), these estimates increase to 81%–84% when using criterion, recall-based tasks with prominent executive demands (Coghill et al. 2014) and a sufficient number of administered trials (Wells et al. 2018).

Applying the construct specificity concerns raised by Sonuga-Barke et al. (2008) and Snyder et al. (2015) to evaluate the tests used to evaluate neuropsychological heterogeneity in ADHD indicates that, to our knowledge, 100% of the available literature either omitted tests of working memory or measured working memory using at least one test with poor construct specificity (Biederman et al. 2004; Coghill et al. 2014; Fair et al. 2012; Geurts et al. 2006; Nigg et al. 2005; Solanto et al. 2001; Sonuga-Barke et al. 2010). As noted above, these studies reported that 30%–37% of ADHD cases are likely to demonstrate working memory deficits based in large part on these less-specific tests. Similarly, 100% of set shifting tasks included in the most recent meta-analysis may be considered non-specific neuropsychological tests rather than specific tests of set shifting (Snyder et al. 2015; Willcutt et al. 2005), and the method used in most studies to derive estimates of inhibitory control from the stop-signal task has been criticized for producing spurious results (Verbruggen et al. 2013).

**Task impurity** Further complicating attempts to identify causal neurocognitive processes in ADHD is the ‘task impurity problem’ (Conway et al. 2005). That is, no task is process pure (Shipstead et al. 2010): all tasks require multiple executive and non-executive neurocognitive abilities for successful performance, and conclusions regarding effect specificity are limited when these correlated but distinct abilities are not measured and simultaneously controlled (Miyake and Friedman 2012). In most cases, the majority of variance in any single test is attributable to processes other than the executive function of interest; multiple tests per construct are critical to isolating EF-specific performance (Willoughby et al. 2016). As noted by Coghill et al. (2014), no study of ADHD-related heterogeneity has simultaneously measured and controlled for all three primary executive functions (Miyake et al. 2000), and conclusions may be limited by the use of single tasks to assess neurocognitive construct(s) (Conway et al. 2005).

## Current Study

Taken together, the literature suggests substantial neurocognitive heterogeneity in pediatric ADHD, but conclusions may be limited because no study to date has simultaneously measured all three primary executive functions (Miyake et al. 2000). In addition, the evidence base has been criticized for using non-specific tests that in many cases were not developed to assess a specific executive function but rather intended to detect gross neurological impairment (Sonuga-Barke et al. 2008). The current study addresses these issues, and is the first to simultaneously and comprehensively assess heterogeneity in executive dysfunction among children with carefully phenotyped ADHD using a counterbalanced battery that includes multiple criterion tests per construct (Coghill et al. 2014). Formative indicators of working memory, inhibitory control, and set shifting were derived, and these participant-level component scores were subjected to reliable change analyses (Jacobson and Truax 1991) that explicitly account for measurement error to objectively define impairment (Kofler et al. 2016; Sarver et al. 2015). We hypothesized that these methodological refinements would produce higher estimates of executive dysfunction than prior studies, such that a majority of children with ADHD would be correctly classified as impaired on at least one of the three primary executive functions. Consistent with models emphasizing causal neurocognitive heterogeneity, we expected to detect a subset of children with ADHD without executive function impairments.

## Method

### Participants

The sample included 136 children aged 8 to 13 years ( $M = 10.37$ ,  $SD = 1.39$ ; 72 boys, 64 girls) from the Southeastern United States, consecutively recruited by or referred to a university-based Children's Learning Clinic (CLC) through community resources (e.g., pediatricians, community mental health clinics, school system personnel, self-referral) between 2015 and 2017. The CLC is a research-practitioner training clinic known to the surrounding community for conducting developmental and clinical child research and providing *pro bono* comprehensive diagnostic and psychoeducational services. Its client base consists of children with suspected learning, behavioral or emotional problems, as well as typically developing children (those without a suspected psychological disorder) whose parents agreed to have them participate in developmental/clinical research studies. All parents and children gave informed consent/assent, and the Florida State University Institutional Review Board approved the study prior to and throughout data collection. Sample ethnicity was mixed with 81 Caucasian Non-Hispanic (62%), 16 African American (12%), 14 Hispanic (11%), 13 multiracial children (10%), and 7 Asian (5%) children.

### Group Assignment

All children with ADHD and their parents completed a comprehensive psychoeducational and diagnostic evaluation that included 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 (2013 Update) allows differential diagnosis according to symptom onset, course, duration, quantity, severity, and impairment in children and adolescents based on DSM-5 criteria (APA 2013), and was supplemented with parent and teacher ratings from the Behavior Assessment System for Children (BASC-2/3; Reynolds and Kamphaus 2015) and ADHD Rating Scale-4/5 (ADHD-4/5; DuPaul et al. 2016). A psychoeducational report was provided to parents.

Fifty-five children met all of the following criteria and were included in the ADHD group ( $n = 55$ ; 38% girls): (1) DSM-5 diagnosis of ADHD Combined ( $n = 39$ ), Inattentive ( $n = 14$ ), or Hyperactive/Impulsive Presentation ( $n = 2$ ) by the directing clinical psychologist based on K-SADS; (2) borderline/clinical elevations on at least one parent and one teacher ADHD subscale; and (3) current impairment based on parent report. All ADHD subtypes/presentations were eligible given the

instability of ADHD subtypes (Valo and Tannock 2010). Psychostimulants ( $n_{prescribed} = 17$ ) were withheld  $\geq 24$  h for testing. To improve generalizability, children with comorbidities were included. Comorbidities reflect clinical consensus best estimates (Kosten & Rounsaville, 1992), and included anxiety (22%), oppositional defiant (13%),<sup>1</sup> depressive (9%), and autism spectrum disorders (5%). Positive screens for reading (9%) and math disability (13%) were defined based on score(s)  $> 1.5$   $SD$  below age-norms on one or more KTEA-3 Academic Skills Battery reading and math subtests, as specified in DSM-5 (APA, 2013).

The Non-ADHD group comprised 81 consecutive case-control referrals (43 girls) who did not meet ADHD criteria, and included both neurotypical children and children with psychiatric disorders other than ADHD. Neurotypical children (77%) had normal developmental histories and nonclinical parent/teacher ratings and were recruited through community resources. Clinically referred and evaluated children who did not meet ADHD criteria were also included in the Non-ADHD group. These Non-ADHD disorders were included to control for comorbidities in the ADHD group, and included best estimate diagnoses of anxiety (13%), autism spectrum (5%), depressive (4%), and oppositional defiant disorders (1%).<sup>1</sup> None of the clinically-evaluated Non-ADHD cases screened positive for learning disorders in reading or math. Importantly, the ADHD and Non-ADHD groups did not differ significantly in the proportion of children diagnosed with a clinical disorder other than ADHD (anxiety:  $BF_{01} = 2.81$ , depression:  $BF_{01} = 4.23$ , ODD:  $BF_{01} = 0.66$ , ASD:  $BF_{01} = 9.57$ , SLD reading:  $BF_{01} = 1.38$ , SLD math  $BF_{01} = 0.41$ ). The Bayes Factor  $BF_{01}$  is an odds ratio indicating support for the null hypothesis that the groups are equivalent ( $H_0$ ) relative to the alternative hypothesis that the groups differ ( $H_1$ ) (see *Bayesian Analyses* section below).

The first 41 Non-ADHD participants underwent an identical evaluation as the ADHD group. Due to funding constraints, the final 40 Non-ADHD participants completed an abbreviated screening evaluation that included parent BASC-3, a 1-subtest IQ screener, and detailed developmental, medical, educational, and psychiatric histories. Neurotypical children did not differ significantly based on whether they received a full or abbreviated evaluation in terms of IQ, gender, age, or BASC hyperactivity T-scores (all  $BF_{01} > 1.37$ ). The abbreviated subgroup had, on average, slightly lower BASC inattention T-scores ( $M =$

<sup>1</sup> As recommended in the K-SADS, oppositional defiant disorder was diagnosed clinically only with evidence of multi-informant/multi-setting symptoms. ODD comorbidity is 48% in the ADHD group and 16% in the Non-ADHD group based on parent-reported symptom counts.



48.0 vs. 53.5,  $BF_{10} = 4.71$ ) and SES ( $M = 45.5$  vs. 53.7,  $BF_{10} = 5.56$ ).

Children were excluded if they presented with gross neurological, sensory, or motor impairment; history of seizure disorder, psychosis, or intellectual disability; or non-stimulant medications that could not be withheld for testing. Additional exclusion criteria were added *a priori* for the abbreviated evaluation subgroup because we were unable to clinically evaluate these cases: previous diagnosis of ADHD or other psychiatric disorders, or BASC-3 inattention/hyperactivity T-scores  $>1.5$  SD above the normative sample mean for age and gender.

## Procedure

Neurocognitive testing occurred as part of a larger battery that involved 1–2 sessions of approximately 3 h each. All tasks were counterbalanced to minimize order effects. Children received brief breaks after each task, and preset longer breaks every 2–3 tasks to minimize fatigue. Children were seated in a caster-wheel swivel chair. Performance was monitored at all times by the examiner, who was stationed just outside of the testing room (out of the child's view) to provide a structured setting while minimizing performance improvements associated with examiner demand characteristics (Gomez and Sanson 1994).

## Measures

### Working Memory Tasks

**Working Memory Reordering** The Rapport et al. (2009) computerized phonological and visuospatial working memory tasks correctly classify children with vs. without ADHD at similar rates as parent and teacher ADHD rating scales (Tarle et al. 2017), and predict hyperactivity (Rapport et al. 2009), attention (Kofler et al. 2010), impulsivity (Raiker et al. 2012), and ADHD-related functional impairments (Friedman et al. 2017; Kofler et al. 2011, 2016). Reliability and validity evidence includes internal consistency ( $\alpha = 0.82$ – $0.97$ ; Kofler et al. 2017), 1- to 3-week test-retest reliability (0.76–0.90; Sarver et al. 2015), and expected magnitude relations with criterion working memory complex span ( $r = 0.69$ ) and updating tasks ( $r = 0.61$ ) (Wells et al. 2018). Internal consistency in the current sample was 0.81 (phonological) and 0.87 (visuospatial).

Both tasks involve serial reordering of characters presented (numbers, black dot locations), and reordering of a target stimulus (letter, red dot location) into the final serial position recalled. Six trials were administered at each set size for each task (3–6 stimuli/trial; 1 stimuli/s). The 24 total trials per task were randomized, then grouped into 2 blocks of 12 trials each,

with short breaks between each block (approximately 1 min) (Kofler et al. 2016). Five practice trials were administered before each task (80% correct required). The phonological task involved mentally reordering and verbally recalling a jumbled series of sequentially presented numbers and letters (e.g., 4H62 is correctly recalled as 246H). The visuospatial task involved mentally reordering a sequentially presented series of spatial locations based on what color dot appeared in each location (black dots in serial order, red dot last) and responding on a modified keyboard. Partial-credit unit scoring (stimuli correct per trial) at each set size (3–6) was used as recommended (Conway et al. 2005).

**Letter Updating** The Miyake et al. (2000) letter memory test was adapted for use with children and exemplifies working memory updating based on the Miyake et al. (2000) model. Working memory updating tasks involve the constant monitoring and rapid addition/deletion of working memory contents (Miyake and Friedman 2012). Similar updating tasks have produced large magnitude ADHD/Non-ADHD between group differences (e.g., Friedman et al. 2017; Raiker et al. 2012). In this computerized task, letters were presented on the screen one at a time, and children were instructed to keep track of the last three letters presented. To ensure the task required continuous updating, children were instructed to rehearse out loud the last three letters by mentally adding the most recent letter and dropping the fourth letter back and then saying the new string of three letters out loud (Miyake et al. 2000). The number of letters presented (4–8 stimuli presented/trial, 1200 ms presentation, 2400 ms ISI) was varied randomly across trials to ensure that successful performance required continuous updating until the end of each trial. A practice block was administered; children advanced to the test phase following three correct trials. Four blocks of three trials each were administered. Children responded via mouse click. The dependent variables were mean stimuli correct per trial recalled in the correct serial order during each of the four task blocks. Internal consistency in the current sample was  $\alpha = 0.75$ .

### Inhibitory Control

**Stop-signal** Task and administration instructions were identical to Alderson et al. (2008). Psychometric evidence includes high internal consistency ( $\alpha = 0.83$ – $0.89$ ), 3-week test-retest reliability (0.72), and convergent validity with other inhibition tests (Soreni et al. 2009). Go-stimuli were displayed for 1000 ms as uppercase letters X and O positioned in the center of a computer screen (500 ms interstimulus interval; total trial duration = 1500 ms). Xs and Os appeared with equal frequency. A 1000 Hz auditory tone (stop-stimulus) was presented randomly on 25% of trials. Stop-signal delay (SSD) – the latency between go- and stop-stimuli presentation – was initially set at 250 ms, and dynamically adjusted  $\pm 50$  ms contingent on performance.

The algorithm was designed to approximate successful inhibition on 50% of stop-trials. In the current study, inhibition success was 60.5%, 58.8%, 59.9%, and 57.6% across the four experimental blocks. Children used a modified response pad to complete two practice and four consecutive experimental blocks of 32 trials/block (8 stop-trials per block). SSD at each of the four blocks is the most direct measure of inhibition in stop-signal tasks that use dynamic stop-signal delays, because SSD changes systematically according to inhibitory success or failure (Alderson et al. 2007; Lijffijt et al. 2005).<sup>2</sup> Internal consistency in the current sample was  $\alpha = 0.80$ .

**Go/no-go** The go/no-go is a response inhibition task in which a motor response must be executed or inhibited based on a stimulus cue (Bezdjian et al. 2009). Children were presented a randomized series of vertical (go stimuli) and horizontal (no-go stimuli) rectangles in the center of a computer monitor (2000 ms presentation, jittered 800–2000 ms ISI to minimize anticipatory responding). They were instructed to quickly click a mouse button each time a vertical rectangle appeared, but to avoid clicking the button when a horizontal rectangle appeared. A ratio of 80:20 go:no-go stimuli was selected to maximize prepotency (Kane and Engle 2003; Unsworth and Engle 2007). Children completed a 10-trial practice (80% correct required) followed by 4 continuous blocks of 25 trials each. Commission errors reflect failed inhibitions (i.e., incorrectly responding to no-go trials), and served as the primary index of inhibitory control during each of the four task blocks. Internal consistency in the current sample was  $\alpha = 0.95$ .

## Set Shifting

**Global-Local** The Miyake et al. (2000) local-global task was adapted for use with children. This computerized task uses Navon (1977) figures, which feature a “global” shape (e.g., a circle) constructed using smaller, “local” figures (e.g., triangles). Figures were presented one at a time in one of four quadrants in a clockwise rotation on a computer monitor (jittered ISI 800–2000 ms). Children were required to shift their response between global and local features depending on which quadrant the figures appeared (top quadrants: global; bottom quadrants: local). Trials with stimuli in the top left or bottom right quadrants involved set shifting (shift trials) because responses required a different rule than the previous trial; trials with stimuli in the top right or bottom left quadrants did not require shifting because they featured the

same rule as the previous trial (non-shift trials). To minimize memory demands, on-screen cues (“big shape”, “small shapes”) remained on-screen next to each quadrant. Sixty trials were administered following three blocks of 6 to 8 practice trials (100% correct required). Internal consistency in the current sample was  $\alpha = 0.86$  (shift trials) and  $\alpha = 0.90$  (no-shift trials).

Children responded via mouse click. Performance data were recorded separately for ‘shift’ and ‘non-shift’ trials. Trials were divided into 4 consecutive blocks to match the number of outcome variables from the working memory and inhibitory control tasks. Reaction time (RT) data was processed following the steps outlined in Miyake et al. (2000) that winsorized the most extreme 2.2% of individual reaction times. First, all individual trial RTs greater than 9500 ms were winsorized to 9500 ms. Second, individual trial RTs greater than 3 standard deviations from each child’s mean RT were winsorized relative to that child’s within-task RT distribution. Shift costs for both response time (speed) and accuracy were computed (van der Ven et al. 2013), calculated separately for each task condition for each child (Speed shift cost =  $RT_{\text{shift}} - RT_{\text{non-shift}}$ ; Accuracy shift cost =  $\% \text{Errors}_{\text{shift}} - \% \text{Errors}_{\text{non-shift}}$ ).

**Number-color** The Miyake et al. (2000) number-letter task was adapted for use with children. A pair of single-digit numbers appeared on the screen, and children were instructed to click either the larger or smaller value depending on the font color (blue = bigger, yellow = smaller; colors selected for maximal discrimination across individuals with all types of color vision). Both digits were the same color on any given trial. To minimize memory demands, on-screen instructions (blue bigger, yellow smaller) remained visible throughout the task. Trials were presented in a semi-random sequence to require shifting every other trial, with an equal number of bigger-smaller and smaller-bigger shifts. RT and accuracy data were recorded separately for ‘shift’ and ‘non-shift’ trials, and processed identically to the global-local data described above. Following an 8-trial practice block (100% correct required), children completed 4 consecutive blocks of 30 trials each (120 total trials; jittered ISI 80–200 ms). Internal consistency in the current sample was  $\alpha = 0.95$  (shift trials) and  $\alpha = 0.87$  (no-shift trials).

Due to a programming error, the first 44 participants (ADHD = 17, Non-ADHD = 27) completed a 60-trial version of the Number-Color task. These data were retained because, accounting for ADHD/Non-ADHD status, shift costs were equivalent for children completing the abbreviated versus full task for both response times ( $BF_{01} = 5.90$ ) and percent correct responses ( $BF_{01} = 7.76$ ), and task version did not interact significantly with ADHD status or task block (all  $BF_{01} > 1.14$ ).

<sup>2</sup> Stop-signal reaction time (SSRT) was also computed for each task block due to current debate in the literature regarding the optimal metric for estimating inhibitory control from the stop-signal task. When substituted for SSD, these SSRT variables failed to load with the inhibitory control variables from the go/no-go task when factor analyzed (loadings = 0.06–0.28), and were therefore excluded from further analysis.

## Executive Function Dimension Reduction

Statistically, we controlled for task impurity by computing Bartlett maximum likelihood weighted averages based on the intercorrelations among task performance scores (DiStefano et al. 2009). Conceptually, this process isolates reliable variance across estimates of each executive function by removing task-specific demands associated with non-executive processes, time-on-task effects via inclusion of four blocks per task, and non-construct variance attributable to other measured executive and non-executive processes (e.g., short-term memory load). Thus, the 36 task performance variables (Supplementary Table 1) were reduced to three principal component estimates (30.06% of variance explained; Supplementary Table 2). A three-component solution was specified *a priori* to derive separate estimates of working memory, inhibitory control, and set shifting based on theory and previous empirical work (e.g., Miyake et al. 2000). The ratio of participants (136) to factors (3) was deemed acceptable (Hogarty et al. 2005). Executive function task data were represented as formative rather than reflective (confirmatory) indicators as recommended (Willoughby et al. 2016). Orthogonal components were specified to maximally control for task impurity (Kofler et al. 2016). By design, the intercorrelations among the varimax-rotated working memory, inhibitory control, and set shifting components were  $r_{all} = 0.00$  ( $p > 0.99$ ). These three executive function component scores (z-scores) were used in all analyses below. Higher scores reflect better working memory and inhibition but worse set shifting.

## Intellectual Functioning (IQ)

All children were administered the WISC-V Short Form (Sattler et al. 2016) or WISC-V Matrix Reasoning subtest (Wechsler 2014) to obtain an estimate of intellectual functioning.

## Socioeconomic Status (SES)

Hollingshead (1975) SES was estimated based on caregiver(s)' education and occupation.

## Objectively-Defined Impairment

Following Sarver et al. (2015) and Kofler et al. (2016), impairment was objectively defined by applying the Jacobson and Truax (1991) model of reliable change to each child's executive function component scores. This method was selected over static cut points (e.g., 10th percentile of Non-ADHD group) because it improves precision by explicitly accounting for measurement unreliability (Jacobson and Truax 1991). Children were classified as Impaired or Not

Impaired in each executive function domain based on whether their score was reliably below the Non-ADHD sample (i.e., difference exceeded chance at  $p < 0.05$ ). This classification was based on computation of the Reliable Change Index (RCI), or the ratio of the difference between the child's score and the Non-ADHD group's mean divided by standard error (computed using each measure's reported test-retest reliability and the *SD* of the Non-ADHD sample; Rule B; Jacobson and Truax 1991) individually for each child for each of the three executive functioning domains. Reported test-retest reliability across all tests was 0.72 to 0.83. The RCI is tested against the  $z$  distribution; impairment is defined as a score that is significantly worse than the Non-ADHD mean given the Non-ADHD group's *SD* and the test's reported reliability. A classification of Impaired indicates that the child is statistically more likely to come from the dysfunctional/impaired population than the functional/neurotypical population (Jacobson and Truax 1991).

Inspection of the RCI data indicated that the impairment cut-offs centered around 1.5 *SD* below the normative sample mean across measures; statistical significance was obtained at different cut points across measures dependent on each measure's test-retest reliability (i.e., for tests with lower reliability, scores further from the mean were required to conclude with  $p < 0.05$  certainty that the child's score was more likely to come from the dysfunctional/impaired population than the functional population).

## Bayesian Analyses

Frequentist statistics were supplemented with Bayesian methods as recommended (Rouder and Morey 2012; Wagenmakers et al. 2016); for our purposes, Bayesian analyses were added because they allow stronger conclusions by estimating the magnitude of support for both the alternative and null hypotheses (Rouder and Morey 2012). JZS default prior scales were used (Rouder and Morey 2012; Wagenmakers et al. 2016). Analyses were conducted using JASP 0.8.2 (JASP Team 2017). Instead of a  $p$ -value, these analyses provide  $BF_{10}$ , which is the Bayes Factor of the alternative hypothesis ( $H_1$ ) against the null hypothesis ( $H_0$ ).  $BF_{10}$  is an odds ratio, where values above 3.0 are considered moderate evidence supporting the alternative hypothesis (i.e., statistically significant evidence for the alternative hypothesis).  $BF_{10}$  values above 10.0 are considered strong ( $>30$  = very strong,  $>100$  = decisive support; Wagenmakers et al. 2016).

Conversely,  $BF_{01}$  is the Bayes Factor of the null hypothesis ( $H_0$ ) against the alternative hypothesis ( $H_1$ ).  $BF_{01}$  is the inverse of  $BF_{10}$  (i.e.,  $BF_{01} = 1/BF_{10}$ ), and is reported when the evidence indicates a lack of an effect (i.e., favors the null hypothesis; Rouder and Morey 2012).  $BF_{01}$  values are interpreted identically to  $BF_{10}$  ( $>3.0$  = moderate,  $>10.0$  = strong,  $>100$  = decisive/extreme support for the null

hypothesis that the ADHD and Non-ADHD groups are equivalent on an outcome; Rouder and Morey 2012).

## Data Analysis Overview

The analytic plan was executed in three tiers. The first Tier examined executive functioning heterogeneity in ADHD by quantifying the proportion of children with ADHD who exhibited impairments in each executive function relative to the local normative comparison group. In the second Tier, we compared demographic and behavioral indicators of children defined as Impaired vs. Not Impaired on each executive function. Finally, we conducted sensitivity analyses in Tier 3 to probe the extent to which our results were influenced by our impairment definition, sample demographics, and data reduction methods.

## Results

### Preliminary Analyses

Outliers beyond 3 *SD* were winsorized relative to the within-group distribution (ADHD, Non-ADHD). This process affected 0.6% (ADHD group) to 0.8% (Non-ADHD group) of data points. Missing data rates were low (0.5% ADHD, 0.2% Non-ADHD) and imputed using the SPSS expectation-maximization function based on all available data because Little's MCAR test indicated that these data were missing completely at random ( $\chi^2 [327] = 108.24, p > 0.99$ ). All parent and teacher ADHD rating scale scores were higher for the ADHD relative to Non-ADHD group as expected (Table 1). The ADHD group demonstrated group-level impairments in working memory ( $d = 1.41$ ;  $BF_{10} = 1.64 \times 10^{10}$ ), inhibitory control ( $d = 0.60$ ;  $BF_{10} = 33.51$ ), and set shifting ( $d = 0.46$ ;  $BF_{10} = 4.29$ ) relative to the Non-ADHD group. In contrast, there was no significant evidence to indicate between-group differences in age ( $BF_{01} = 2.34$ ), gender ( $BF_{01} = 1.10$ ), ethnicity ( $BF_{01} = 2.29$ ), IQ ( $BF_{10} = 2.00$ ), or SES ( $BF_{01} = 4.48$ ); we therefore report simple model results with no covariates.

### Tier 1: Executive Functioning Impairments in ADHD

As shown in Fig. 1, 89% of children with ADHD were classified as Impaired on at least one executive function component ( $n = 49$  of 55). This finding corresponded to a median log odds ratio ( $OR_{log}$ ) of 3.39, with diagnostic sensitivity of 89.1% and specificity of 80.3% (Supplementary Table 3). Within specific executive function domains, 62% of children with ADHD were classified as Impaired in working memory ( $OR_{log} = 2.62$ ), 27% Impaired in inhibitory control ( $OR_{log} = 1.49$ ), and 38% Impaired in set shifting ( $OR_{log} = 2.36$ ). Approximately half of the children with ADHD showed

impairment on a single executive function (54%), and an additional third showed impairments on two (31%) or all three (4%) executive functions.

Only 11% (6 of 55) of children with ADHD were classified as Not Impaired on all three executive functions. Following Coghill et al. (2014), we inspected these cases to determine whether they 'just missed' the threshold for definition of a deficit. Qualitative inspection of these children's data suggest that 5 of the 6 Not Impaired children fell just above the impairment cut-off on one of the three executive function components (i.e., they had scores at/below the 20th percentile relative to the Non-ADHD group;  $n = 4$  for working memory,  $n = 1$  for inhibitory control); the final Not Impaired child with ADHD showed below average performance on two of the three executive function components (27th–30th percentile for inhibitory control and set shifting) but above average working memory (75th percentile). Formal statistical tests could not be conducted due to low cell counts.

### Tier 2: Profiles of Impaired vs. Not Impaired Children

Exploratory analyses were conducted to probe for demographic and behavioral predictors as a function of impairment status, separately for each of the three executive function components. Results are based on the  $n = 96$  cases with teacher ratings, and should be considered tentative given the relatively small subgroup sample sizes (Supplementary Table 4).

#### Demographic Characteristics

There was no significant evidence for, and in most cases significant evidence *against*, differences in age, gender, IQ, SES, or medication status as a function of impairment status on each executive function (all  $BF_{10} < 1.26$ ) with one exception: Children with working memory impairment ( $M_{age} = 9.91$ ,  $SD = 1.41$ ) were slightly younger than children without working memory impairment ( $M_{age} = 10.82$ ,  $SD = 1.45$ ) ( $BF_{10} = 11.35$ ,  $d = 0.64$ ).

#### ADHD Symptom Severity

ADHD symptom profiles associated with impairment in each executive function domain were assessed using a series of 2 (EF Impairment: no/yes)  $\times$  2 (informant: parent, teacher)  $\times$  2 (ADHD symptom domain: inattention, hyperactivity/impulsivity) Bayesian mixed-model ANOVAs. Informant and symptom domain were treated as within-subject factors to maximize power by including multiple indices per construct and minimizing the number of omnibus models. Main effects are corrected for multiple testing by fixing to 0.5 the prior probability that the null hypothesis holds across all comparisons (JASP Team 2017; Westfall, Johnson, & Utts, 1997). Separate models were run for each of the three executive



**Table 1** Sample and demographic variables

Variable	ADHD (N = 55)		Non-ADHD (N = 81)		Cohen's d	BF <sub>10</sub>	BF <sub>01</sub>	p
	M	SD	M	SD				
Gender (Boys/Girls)	34/21		38/43		–		1.10	0.09, <i>ns</i>
Ethnicity (AA/A/C/H/M)	9/0/39/4/3		9/7/45/10/10		–		2.29	0.06, <i>ns</i>
Age	10.18	1.40	10.51	1.37	–0.24		2.34	0.18, <i>ns</i>
SES	46.92	12.38	48.27	12.35	–0.11		4.48	0.53, <i>ns</i>
FSIQ	102.75	14.65	108.02	12.06	–0.40	2.00		0.02
BASC-2/3 Attention Problems (T-score)								
Parent	66.07	8.16	53.94	10.79	1.24	1.02 × 10 <sup>8</sup>		<0.001
Teacher	63.11	9.05	54.88	11.35	0.82	154.99		<0.001
BASC-2 Hyperactivity (T-score)								
Parent	67.31	12.30	53.60	10.30	1.23	8.30 × 10 <sup>7</sup>		<0.001
Teacher	61.91	14.85	53.93	12.88	0.57	5.76		0.007
Executive Function Component Scores (Z-scores)								
Working Memory	–0.69	0.95	0.47	0.73	–1.41	1.64 × 10 <sup>10</sup>		<0.001
Inhibitory Control	–0.34	1.09	0.23	0.87	–0.60	33.51		<0.001
Set Shifting	–0.27	1.32	0.18	0.65	–0.46	4.29		0.009

*P*-values are not corrected for family-wise error, and are included for illustrative purposes to allow interested readers to compare Bayesian and frequentist results. BF<sub>10</sub> = Bayes Factor for the alternative hypothesis over the null hypothesis (values ≥ 3.0 indicate significant between-group differences). BF<sub>01</sub> = Bayes Factor for the null hypothesis over the alternative hypothesis (values ≥ 3.0 indicate significant between-group equivalence; BF<sub>01</sub> = 1/BF<sub>10</sub>). BASC behavior assessment system for children. Teacher BASC data were not collected for the *n* = 40 Non-ADHD children who received the abbreviated screening evaluation (Non-ADHD *n* = 41 for BASC teacher comparisons). Ethnicity: AA african american, A asian, C caucasian non-hispanic, H hispanic, M multiracial

function domains (working memory, inhibitory control, set shifting). In all three models there was significant evidence for a main effect of informant (parent > teacher; BF<sub>10</sub> = 303.50), but significant evidence *against* effects of symptom domain (inattention = hyperactivity/impulsivity; BF<sub>01</sub> = 6.45) and informant × domain interactions (BF<sub>01</sub> > 6.67). The pattern and interpretation of results were unchanged when controlling for age, gender, SES, IQ, and medication status; and when examining just the ADHD sample. Cell sizes were insufficient to compare Impaired vs. Not Impaired Non-ADHD children (i.e., only 4–8 impaired Non-ADHD children per executive function as expected due to our method of defining impairment based on the Non-ADHD distribution).

**Working Memory** Results indicated significantly higher ADHD symptoms for children with impaired working memory (*n* = 37) relative to children with unimpaired working memory (*n* = 59) (BF<sub>10</sub> = 5.23 × 10<sup>4</sup>, *d* = 0.36). Relative to the main effects model, there was significant evidence *against* all interaction effects, indicating that the impaired working memory subgroup showed higher ADHD symptoms across domains and informants (all BF<sub>01</sub> > 4.03; please see Figure accompanying Supplementary Table 4).

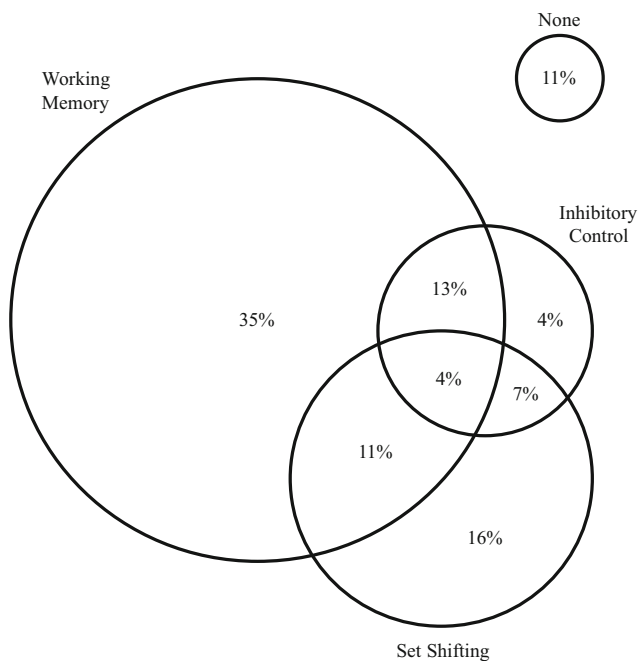
**Inhibitory Control** Results indicated significant evidence *against* ADHD symptom differences for children with

impaired inhibitory control (*n* = 19) relative to children with unimpaired inhibitory control (*n* = 77) (BF<sub>01</sub> = 5.80, *d* = 0.04). Relative to the main effects model, there was significant evidence *against* the inhibition deficit × symptom domain interaction (BF<sub>01</sub> = 6.31); results were inconclusive for the inhibition × informant interaction (BF<sub>01</sub> = 1.64) and the 3-way interaction (BF<sub>01</sub> = 2.19).

**Set Shifting** Results indicated significant evidence *against* ADHD symptom differences for children with impaired set shifting (*n* = 24) relative to children with unimpaired set shifting (*n* = 72) (BF<sub>01</sub> = 7.34, *d* = 0.02). Relative to the main effects model, there was significant evidence *against* all interactions (BF<sub>01</sub> > 3.09).

### Tier 3: Sensitivity Analyses

Finally, we probed the extent to which results were influenced by our use of the reliable change method for defining impairment (Jacobson and Truax 1991), sample demographics, and specification of orthogonal components. Results based on the reliable change method reported above were highly consistent with results obtained by applying the static 10th percentile cutoff used in several previous studies of ADHD-related neurocognitive heterogeneity. That is, defining impairment as the 10th percentile of the Non-ADHD group resulted in 91% of



**Fig. 1** Visual heuristic showing the proportion of children with ADHD classified as Impaired on each executive function based on the Jacobson and Truax (1991) model of reliable change. Circle sizes are proportionate to the percentage of children identified as Impaired in each domain. Values do not sum to 100% due to rounding; precise values are given in Supplementary Tables 3a-d

ADHD children being classified as Impaired on at least one executive function ( $n = 50$  of 55; sensitivity = 90.91% [80.05, 96.98], specificity = 75.31% [64.47, 84.22]). Highly similar results were also obtained when controlling for age, gender, and SES by residualizing for these variables and then re-computing impairment estimates (82% of ADHD sample classified as impaired; sensitivity = 81.82% [69.10, 90.92], specificity = 80.25% [69.91, 88.27]). Highly similar results were also obtained when allowing covariation among the three executive functioning estimates by specifying an oblique rather than orthogonal factor solution (84% of ADHD sample classified as impaired; sensitivity = 83.64% [71.20, 92.23], specificity = 81.48% [73.30, 89.25]). Finally, we tested a 1-factor solution to probe the extent to which results were influenced by modeling separate executive function components. This overall estimate of executive functioning showed high specificity (98.77% [93.31, 99.97]) but poor sensitivity (34.55% [22.24, 48.58]), supporting ADHD-related heterogeneity across executive function domains.

## Discussion

The current study was the first to comprehensively assess executive functioning heterogeneity in pediatric ADHD, using a carefully phenotyped sample of children with ADHD (Coghill et al. 2014), multiple criterion tests of

each primary executive function (working memory, inhibitory control, set shifting; Miyake et al. 2000), and objectively-defined impairment (Jacobson and Truax 1991). Results indicate that 89% of children with ADHD demonstrated executive functioning deficit(s), with sensitivity (89%) and specificity (80%) estimates that are similar to or exceed the diagnostic utility of ADHD symptom checklists (Tarle et al. 2017). Together, these estimates are considerably higher than previous reports and, if replicated, suggest that cognitively-informed test batteries may allow for the identification of a relatively small number of causal pathways to the ADHD phenotype (Coghill et al. 2005). Stated differently, the current findings confirm that neurocognitive heterogeneity is a valid phenomenon in ADHD (only 35% were impaired when defining executive function as a unitary, 1-component construct), but that this heterogeneity may be more circumscribed than previously suggested with 89% showing impairment in at least one of the three primary executive functions. The detection of neuropsychologically impaired subtypes (Nigg et al. 2005) has the potential to inform etiological models of ADHD, identify novel intervention targets, and tailor interventions to maximize efficacy (Chacko et al. 2014).

Approximately 10% of children with ADHD were classified as unimpaired on all three executive functions. We hypothesized that there would be an ADHD subgroup with intact executive functioning, and these results appear to support this prediction. Several additional neurocognitive processes have been proposed as candidate causal mechanisms in ADHD, including altered reinforcement gradients and delay aversion, temporal processing (timing) deficits, response inconsistency/variability, hypo-arousal and/or -activation, and slowed processing speed (Fair et al. 2012; Sonuga-Barke et al. 2008). Although process analysis suggests that poor performance on tests of delay aversion (Patros et al. 2015), response variability (Kofler et al. 2014), and emotion dysregulation (van Cauwenberge et al. 2015) by children with ADHD may be parsimoniously explained by associations with executive dysfunction, these abilities were not assessed in the current study and merit consideration in future work. Alternatively, inspection of subject-level data from the six incorrectly-classified children with ADHD suggested below average performance that in most cases 'just missed' our objectively-defined cut-offs (Coghill et al. 2014). Future work is needed to determine the threshold below which underdeveloped executive function(s) produce that hallmark ADHD phenotype, whether individual symptoms are associated with specific EF deficits, and whether children with ADHD show differential functional impairments based on their executive function profile. Biederman et al. (2004), for example, showed that executive dysfunction predicted significantly lower academic achievement, and Kofler et al. (2016) demonstrated differential associations across functional impairment domains.

A small group of 4–8 Non-ADHD children were classified as impaired on each executive function. Although this result was expected, and in large part a statistical artifact of defining impairment based on the Non-ADHD distribution, it highlights the need to consider trait variation in the typically developing population (Fair et al. 2012) when developing and refining causal models of ADHD (Coghill et al. 2005). In the current study, there were too few misclassified Non-ADHD children to conduct inferential statistics. Future work is needed to determine whether Non-ADHD children with executive dysfunction show elevated but subclinical ADHD symptoms, whether these children reflect clinical control cases, and/or whether these children possess protective personal and social assets (Lerner et al. 2009) that buffer against ADHD despite this risk factor.

A key assumption underlying the current study's methods and conclusions is that executive function deficits are causal mechanisms responsible, in part, for manifestation of the ADHD phenotype. It is important to note that causality cannot be assumed from the current results, and to acknowledge alternate conceptualizations of executive function deficits in ADHD (e.g., epiphenomenal, third variable influences; van Lieshout et al. 2013). Indeed, although our test battery demonstrates that most children with ADHD have executive functioning deficits, it cannot inform whether these deficits occur due to impaired top-down processing or bottom-up signaling that should have recruited the top-down activity (Sonuga-Barke et al. 2008). Further, the extent to which an executive function deficit is a cause, correlate, or effect may differ from child to child and from executive function to executive function. Proximal environmental influences appear to exert significant influences on state-based ADHD behaviors as a function of their executive function demands (Kofler et al. 2016). For example, evidence from carefully controlled experimental studies suggests that ADHD-related inattentive and hyperactive behavior can be produced and rarefied by manipulating working memory demands (Hudec et al. 2015; Kofler et al. 2010; Rapport et al. 2009), whereas similar manipulations of inhibition demands have failed to systematically affect hyperactive behavior (Alderson et al. 2012). Similarly, large-scale longitudinal studies show that age-related reductions in ADHD symptoms are predicted by improvements in working memory but not inhibitory control (Karalunas et al. 2017). Future work is needed to determine whether set shifting produces similar experimental/longitudinal changes, clarify the causal role of each executive function in ADHD-related behavioral and functional impairments, and determine whether neurologically informed subtypes provide clinical utility for maximizing treatment efficacy via improved targeting (Chacko et al. 2014; Nigg et al. 2005).

## Limitations

The current study was the first to comprehensively assess executive functioning in ADHD using a criterion battery of cognitively-informed tests, objectively-defined impairment criteria that explicitly accounted for measurement unreliability, and a relatively large sample of carefully phenotyped children with and without ADHD. Despite these methodological refinements, the following limitations must be considered when interpreting results. Given that co-occurring conditions are common in ADHD (Wilens et al. 2002), inclusion of children with these comorbidities was important to maximize external validity and generalizability of our findings. We attempted to balance external and internal validity threats by recruiting a Non-ADHD group matched for the number of these Non-ADHD disorders; however, controlling for the number of other disorders does not perfectly equate the groups, and as such future work is needed to determine how more 'pure' ADHD samples compare to non-disordered children – particularly given that neither executive dysfunction nor behavioral symptoms (e.g., difficulty concentrating, restlessness) appear unique to ADHD (e.g., Snyder, 2013; Youngstrom et al. 2010). Independent replications with larger samples, naturalistic outcomes, and a broader sampling of children with other clinical disorders are needed to assess the extent to which the executive dysfunction profiles identified herein provide meaningful prediction toward specific symptoms and ecologically-valid impairment domains.

Children with impaired working memory exhibited moderately higher parent- and teacher-reported ADHD symptoms than children with unimpaired working memory, whereas there was significant evidence *against* such differences as a function of inhibitory control and set shifting deficits. This pattern of results appears consistent with ADHD conceptual models highlighting equifinality, to the extent that the vast majority of the ADHD group demonstrated impairment in at least one EF domain (i.e., multiple combinations of impairments leading to the same level of ADHD symptoms; e.g., Sonuga-Barke et al. 2008). Alternatively, it may reflect restricted range associated with our stringent ADHD criteria, potentially in combination with the relatively small subgroup sample sizes for these analyses.

Importantly, our fractionation of executive functioning into distinct domains of working memory, inhibitory control, and set shifting was consistent with theory and empirical work (Miyake et al. 2000). At the same time, these executive subfunctions are themselves not unitary constructs. For example, working memory can be further fractionated into at least three central executive processes (updating, dual-processing, temporal reordering; the 'working' components of working memory; Rapport et al. 2013) as well as three non-executive short-term memory stores (phonological, visual-spatial, episodic; Baddeley 2012), each of which is functionally and

anatomically distinguishable (Wager and Smith 2003). Similarly, inhibitory control can be fractionated into processes involved in action cancellation and action restraint (Alderson et al. 2007); we are unaware of any studies documenting subcomponents of set shifting. It remains to be seen whether ADHD is associated globally or locally with some/all of these executive subprocesses.

## Clinical and Research Implications

The current results confirm that ADHD is characterized by neurocognitive heterogeneity (Nigg et al. 2005), while suggesting that applying contemporary criteria informed by cognitive science to executive function may provide improved precision for identifying a smaller number of neuropsychologically impaired subtypes than previously described. If replicated, these findings would indicate that executive dysfunction is more prevalent in ADHD than previously estimated, and provide a strong empirical basis for a line of basic and applied research to determine whether classifying children with ADHD based on their executive function profile provides a meaningful mechanism toward understanding functional impairments, predicting outcomes, and improving treatment efficacy via improved targeting and neurocognitively-informed modifications (Chacko et al. 2014).

**Acknowledgements** This work was supported in part by an NIH grant (R34 MH102499-01, PI: Kofler). The sponsor had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

## Compliance with Ethical Standards

**Conflict of Interest** The authors have no conflicts of interest to report.

**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 was obtained from all individual participants included in the study.

## References

- 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.
- Alderson, R. M., Rapport, M. D., Sarver, D. E., & Kofler, M. J. (2008). ADHD and behavioral inhibition: A re-examination of the stop-signal task. *Journal of Abnormal Child Psychology*, *36*, 989–998.
- 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*(5), 487–505.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, D.C.: American Psychiatric Association.
- Baddeley, A. (2007). *Working memory, thought, and action*. Oxford Press.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.
- Bezdzjian, S., Baker, L. A., Lozano, D. I., & Raine, A. (2009). Assessing inattention and impulsivity in children during the Go/NoGo task. *British Journal of Developmental Psychology*, *27*, 365–383.
- Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., et al. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology*, *72*(5), 757.
- Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Reviews Neuroscience*, *3*(8), 617–628.
- Chacko, A., Kofler, M., & Jarrett, M. (2014). Improving outcomes for youth with ADHD: A conceptual framework for combined neurocognitive and skill-based treatment approaches. *Clinical Child and Family Psychology Review*, *17*(4), 368–384.
- Coghill, D., Nigg, J., Rothenberger, A., Sonuga-Barke, E., & Tannock, R. (2005). Whither causal models in the neuroscience of ADHD? *Developmental Science*, *8*(2), 105–114.
- Coghill, D. R., Seth, S., & Matthews, K. (2014). A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: advancing beyond the three-pathway models. *Psychological Medicine*, *44*, 1989–2001.
- 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 and Review*, *12*(5), 769–786.
- Cortese, S., Kelly, C., Chabernaud, C., Proal, E., Di Martino, A., Milham, M. P., & Castellanos, F. X. (2012). Toward systems neuroscience of ADHD: A meta-analysis of 55 fMRI studies. *American Journal of Psychiatry*, *169*(10), 1038–1055.
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation*, *14*, 1–11.
- DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (2016). *ADHD rating scale-5 for children and adolescents: Checklists, norms, and clinical interpretation*. New York: Guilford Press.
- 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*, *125*, 309–331.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation*, *44*, 145–200.
- Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(17), 6769–6774.
- Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204.
- 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.
- Geurts, H. M., van der Oord, S., & Crone, E. A. (2006). Hot and cool aspects of cognitive control in children with ADHD: Decision-



- making and inhibition. *Journal of Abnormal Child Psychology*, 34(6), 813–824.
- Gomez, R., & Sanson, A. (1994). Effects of experimenter and mother presence on the attentional performance and activity of hyperactive boys. *Journal of Abnormal Child Psychology*, 22, 517–529.
- Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality factor solutions in exploratory factor analysis: The influence of sample size, communalities, and overdetermination. *Educational and Psychological Measurement*, 65, 202–226.
- Hollingshead, A.B. (1975). *Four factor index of social status*. Yale: New Haven.
- Hudec, K. L., Alderson, R. M., Patros, C. H. G., Lea, S. E., Tarle, S. J., & Kasper, L. J. (2015). Hyperactivity in boys with attention-deficit/hyperactivity disorder (ADHD): The role of executive and non-executive functions. *Research in Developmental Disabilities*, 45, 103–109.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- 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.
- JASP Team. (2017). JASP (Version 0.8.2 software).
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47–70.
- Karalunas, S. L., Gustafsson, H. C., Dieckmann, N. F., Tipsord, J., Mitchell, S. H., & Nigg, J. T. (2017). Heterogeneity in development of aspects of working memory predicts longitudinal attention deficit hyperactivity disorder symptom change. *Journal of Abnormal Psychology*, 126(6), 774.
- 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., et al. (1997). Schedule for affective disorders and schizophrenia for school-age children (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 980–988.
- 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, 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., Alderson, R. M., Raiker, J. S., Bolden, J., Sarver, D. E., & Rapport, M. D. (2014). Working memory and intraindividual variability as indicators in ADHD. *Neuropsychology*, 28, 459–471.
- Kofler, M. J., Sarver, D. E., Spiegel, J. A., Day, T. N., Harmon, S. L., & Wells, E. L. (2016). Heterogeneity in ADHD: Neurocognitive predictors of peer, family, and academic functioning. *Child Neuropsychology*, 23(6), 733–759.
- Kofler, M. J., Sarver, D. E., Harmon, S. L., Moltisanti, A., Aduen, P. A., Soto, E. F., & Ferretti, N. (2017). Working memory and organizational skills problems in ADHD. *Journal of Child Psychology and Psychiatry*.
- Kofler, M. J., Sarver, D. E., Harmon, S. L., Moltisanti, A., Aduen, P. A., Soto, E. F., & Ferretti, N. (2018). Working memory and organizational skills problems in ADHD. *Journal of Child Psychology and Psychiatry*, 59(1), 57–67.
- Kosten, T. A., & Rounsaville, B. J. (1992). Sensitivity of psychiatric diagnosis based on the best estimate procedure. *The American Journal of Psychiatry*, 149(9), 1225.
- Lerner, J. V., Phelps, E., Forman, Y. E., & Biowers, E. P. (2009). Positive youth development. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 524–558). NJ: Wiley.
- 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), 216–222.
- Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44(4), 377–384.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8–14.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3), 353–383.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, 57(11), 1224–1230.
- Pa, J., Possin, K. L., Wilson, S. M., Quitania, L. C., Kramer, J. H., Boxer, A. L., et al. (2010). Gray matter correlates of set-shifting among neurodegenerative disease, mild cognitive impairment, and healthy older adults. *Journal of the International Neuropsychological Society*, 16(04), 640–650.
- Patros, C. H. G., Alderson, R. M., Lea, S. E., Tarle, S. J., Kasper, L. J., & Hudec, K. L. (2015). Visuospatial working memory underlies choice-impulsivity in boys with ADHD. *Research in Developmental Disabilities*, 38, 134–144.
- Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014). ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *International Journal of Epidemiology*, 43(2), 434–442.
- Raiker, J. S., Rapport, M. D., Kofler, M. J., & Sarver, D. E. (2012). Objectively-measured impulsivity and ADHD: Testing competing predictions from the working memory and behavioral inhibition models of ADHD. *Journal of Abnormal Child Psychology*, 40, 699–713.
- Rapport, M. D., Bolden, J., Kofler, M. J., Sarver, D. E., Raiker, J. S., & Alderson, R. M. (2009). Hyperactivity in boys with ADHD: A ubiquitous core symptom 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. *Clinical Psychology Review*, 33, 1237–1252.
- Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 20(6), 1102–1113.
- Reynolds, C.R., & Kamphaus, R.W. (2015). *BASC-3: Behavior assessment system for children*.
- Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research*, 47(6), 877–903.
- Sarver, D. E., Rapport, M. D., Kofler, M. J., Raiker, J. S., & Friedman, L. M. (2015). Hyperactivity in ADHD: Impairing deficit or compensatory behavior? *Journal of Abnormal Child Psychology*, 43, 1219–1232.

- Sattler, J., Dumont, R., & Coalson, D. (2016). *Assessment of Children: WISC-V and WPPSI-IV*. San Diego: Sattler Press.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, *50*, 245–276.
- Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychological Bulletin*, *139*(1), 81.
- Snyder, H. R., Miyake, A., & Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. *Frontiers in Psychology*, *6*.
- Solanto, M. V., Abikoff, H., Sonuga-Barke, E., Schachar, R., Logan, G. D., Wigal, T., et al. (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. J., Sergeant, J. A., Nigg, J., & Willcutt, E. (2008). Executive dysfunction and delay aversion in attention deficit hyperactivity disorder: nosologic and diagnostic implications. *Child and Adolescent Psychiatric Clinics of North America*, *17*(2), 367–384.
- 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*, 345–355.
- Soreni, N., Crosbie, J., Ickowicz, A., & Schachar, R. (2009). Stop signal and Conners' CPT: Test-retest reliability of two inhibition measures in ADHD children. *Journal of Attention Disorders*, *13*, 137–143.
- St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*(4), 745–759.
- Tarle, S. J., Alderson, R. M., Patros, C. H. G., Lea, S. E., Hudec, K. L., & Arrington, E. F. (2017). Attention-deficit/hyperactivity disorder and phonological working memory: Methodological variability affects clinical and experimental performance metrics. *Neuropsychology*, *31*(4), 383–394.
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132.
- Valo, S., & Tannock, R. (2010). Diagnostic instability of DSM-IV ADHD subtypes: Effects of informant source, instrumentation, and methods for combining symptom reports. *Journal of Clinical Child & Adolescent Psychology*, *39*, 749–760.
- van Cauwenberge, V., Sonuga-Barke, E. J. S., Hoppenbrouwers, K., Van Leeuwen, K., & Wiersma, J. R. (2015). "Turning down the heat": Is poor performance of children with ADHD on tasks tapping "hot" emotional regulation caused by deficits in "cool" executive functions? *Research in Developmental Disabilities*, *47*, 199–207.
- Van der Sluis, S., De Jong, P. F., & Van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *427*–449.
- van der Ven, S. H., Kroesbergen, E. H., Boom, J., & Leseman, P. P. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *British Journal of Developmental Psychology*, *31*(1), 70–87.
- van Lieshout, M., Luman, M., Buitelaar, J., Rommelse, N. N., & Oosterlaan, J. (2013). Does neurocognitive functioning predict future or persistence of ADHD? A systematic review. *Clinical Psychology Review*, *33*(4), 539–560.
- Verbruggen, F., Chambers, C. D., & Logan, G. D. (2013). Fictitious inhibitory differences: how skewness and slowing distort the estimation of stopping latencies. *Psychological Science*, *24*(3), 352–362.
- Wagenmakers, E., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, *25*(3), 169–176.
- Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 255–274.
- Wechsler, D. (2014). *Wechsler Intelligence Scale for Children-Fourth/Fifth Edition*. San Antonio: Pearson.
- Wells, E. L., Kofler, M. J., Soto, E. F., Schaefer, H., & 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.
- Westfall, P. H., Johnson, W. O., & Utts, J. M. (1997). A Bayesian perspective on the Bonferroni adjustment. *Biometrika*, *84*(2), 419–427.
- Wiemers, E. A., & Redick, T. S. (2018). Working memory capacity and intra-individual variability of proactive control. *Acta Psychologica*, *182*, 21–31.
- Wilens, T. E., Biederman, J., Brown, S., Tanguay, S., Monuteaux, M. C., Blake, C., & Spencer, T. J. (2002). Psychiatric comorbidity and functioning in clinically referred preschool children and school-age youths with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 262–268.
- Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology*, *27*, 35–78.
- Willoughby, M. T., Blair, C. B., & The Family Life Project Investigators. (2016). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment*, *28*(3), 319–330.
- Youngstrom, E. A., Arnold, L. E., & Frazier, T. W. (2010). Bipolar and ADHD comorbidity: both artifact and outgrowth of shared mechanisms. *Clinical Psychology: Science and Practice*, *17*(4), 350–359.
- Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology*, *16*(7), 653–667.