



Published in final edited form as:

Clin Psychol Sci. 2017 March ; 5(2): 195–211. doi:10.1177/2167702616668320.

The role of speed in ADHD-related working memory deficits: A time-based resource-sharing and diffusion model account

Alexander Weigard and Cynthia Huang-Pollock

The Pennsylvania State University

Abstract

Several recent commentaries suggest that, for psychological science to move beyond “homuncular” explanations for cognitive control, it is critically important to examine the role of basic and computationally well-defined processes (e.g. cognitive processing speed). Correlational evidence has previously linked slow speed to working memory (WM) deficits in ADHD, but the directionality of this relationship has not been investigated experimentally and the mechanisms through which speed may influence WM are unclear. Herein, we demonstrate in school-aged children with and without ADHD, that manipulating speed (indexed with the diffusion model) within a WM paradigm reduces WM capacity due to an increase in cognitive load, in a manner that is consistent with predictions of the time-based resource-sharing model of WM. Results suggest slow speed is a plausible cause of WM deficits in ADHD, provide a mechanistic account of this relationship, and urge the exploration of non-executive neurocognitive processes in clinical research on etiology.

Keywords

ADHD; working memory; time-based resource-sharing; drift diffusion model

Impairment in working memory (WM), the ability to concurrently process and store information (Baddeley, 2012; Miyake & Shah, 1999; Oberauer, Lewandowsky, Farrell, Jarrold & Greaves, 2012; Unsworth & Engle, 2007), is one of the most consistently identified neurocognitive abnormalities in Attention-Deficit Hyperactivity Disorder (ADHD) (Kasper, Alderson, & Hudec, 2012; Schoechlin & Engle, 2005) and potentially one of the most impairing. WM capacity predicts reading comprehension over and above storage-only tasks (Daneman & Merikle, 1996) and has been linked to domains as diverse as verbal ability (Alloway, Gathercole, Kirkwood, & Elliott, 2009), spatial ability (Miyake, Friedman, Rettinger, Shah & Hegarty, 2001), mathematical competency (Nyroos & Wiklund-Hornqvist, 2012; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007), and a host of other skills critical for academic success (Engle, Kane & Tuholski, 1999; Engle, Tuholski,

Correspondence concerning this article should be addressed to Alexander Weigard, Department of Psychology, Moore Building, Penn State University, University Park, PA, 16802. asw5097@psu.edu.

The authors have no conflicts of interest to report.

Authorship Statement: C.H.-P. and A.W. developed the study design and concept. A.W. constructed the measure, analyzed the data, and interpreted the data under the supervision of C.H.-P. C.H.-P. and A.W. drafted the paper together. Both authors approved the final version of the paper for submission.

Laughlin & Conway, 1999; Gathercole and Pickering, 2000). It is therefore not surprising that WM impairment is associated with the significant academic and classroom behavioral problems commonly observed in children with ADHD (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Alloway, Gathercole, Holmes, Place, Elliott, Hilton, 2009; Gathercole and Pickering, 2000; Gremillion & Martel, 2012; Miller et al., 2013), and that children with ADHD show greater difficulty developing automatic, skilled performance on tasks with high but not low WM demands (Huang-Pollock & Karalunas, 2010).

The conventional view within the scientific community studying ADHD has generally understood WM impairment to be one of several examples of executive function (EF) deficits observed in affected children (Barkley, 1997; Burgess et al., 2010; Castellanos & Tannock, 2002; Coghill et al., 2005; Kofler et al., 2014; Rapport et al., 2008; Willcutt et al., 2005). Indeed, the ability to flexibly switch between effortful processing and the storage of information is related, both conceptually and neurobiologically, to other putative EFs (Engle, Kane & Tuholski, 1999), where EF is defined as a higher order ability linked to the prefrontal cortex that allows individuals to plan, set shift, and override pre-potent responses in favor of goal-directed behavior (Barkley, 1997; Burgess et al., 2010; Castellanos & Tannock, 2002; Coghill et al., 2005; Kofler et al., 2014; Rapport et al., 2008; Willcutt et al., 2005).

This view also broadly aligns with a preeminent theory of individual differences in WM ability from basic cognitive science, which understands the core cognitive process being measured in tasks of WM to be executive attention (Engle, 2002; Unsworth and Engle, 2007). Classical models of WM (Baddeley & Hitch, 1974) posit that a “central executive” attention system manages the storage and rehearsal operations that maintain memory items, while preventing interference from other tasks. And, WM capacity is strongly correlated with performance on tasks indexing the ability to effortfully control attention, such as the Stroop and anti-saccade paradigms (Kane, Bleckley, Conway & Engle, 2001; Kane & Engle, 2002; Kane & Engle, 2003; Meier & Kane, 2013). Thus, the argument is that WM reflects an individual's capacity to use controlled attention, and individual differences in WM reflect individual differences in executive attention (Engle, 2002). Indeed, ADHD status and symptom severity are associated with poor performance on the antisaccade (Carr, Nigg & Henderson, 2006; Goto et al., 2010; Nigg, Butler, Huang-Pollock & Henderson, 2002) and Stroop paradigms (Barkley, Grodzinsky & Dupaul, 1992; Ikeda, Okuzumi & Kokubun, 2013), supporting the idea that WM deficits in ADHD can be attributed to weaknesses in a broad executive attentional construct. This suggestion is explicitly articulated by the functional working memory model of ADHD (Rapport et al., 2008; Kofler et al., 2014), which posits that ADHD-related deficits in a central executive attentional controller are the core cause of both WM impairment and broader manifestations of the disorder, including impulsivity and reaction time (RT) variability.

However, a major limitation of such conceptualizations is that the EF or “central executive” construct is difficult to define in a parsimonious and mechanistic way. Theories of WM and other executive functions have been famously criticized (Monsell & Driver, 2000) for relying on a control “homunculus”, or an intelligent agent who simply carries out the complex operations required by the theory (e.g., shifting behavior towards higher order

goals, supervising WM maintenance) without reference to specific mechanisms. Attempts have been made to fractionate this agent's duties into several more specific processes in the hope that these discrete components of executive control would be more easy to define and study mechanistically (Miyake et al., 2000; Monsell & Driver, 2000). More recently, Verbruggen et al. (2014) argued that the more specific executive processes identified by this work (e.g., "inhibition") are still poorly defined and are fundamentally descriptive, rather than explanatory, constructs. As an alternative approach, these authors and others have called for the investigation of how basic, easily measurable and computationally well-defined processes may underlie cognitive control (Vandierendonck, 2016; Verbruggen et al., 2014). If the contribution of these basic processes to individual differences in higher order EFs, such as WM, is clarified, better-specified models of EF and EF deficits in clinical populations can be developed and tested. Indeed, studies from the emerging field of model-based cognitive neuroscience have demonstrated the power of mathematically-specified theories for linking psychological explanations to neural mechanisms (Forstmann & Wagenmakers, 2014), highlighting the need for formalized theories of clinical dysfunction.

Processing speed as a basic underpinning of WM capacity in ADHD

In an example directly relevant to the case of WM deficits in childhood ADHD, significant evidence has accumulated over the past two decades that the basic speed of information processing, a clearly defined and computationally simple construct, may be a key determinant of children's WM capacity. Processing speed refers to the general efficiency (typically operationalized by RT) with which a given individual can complete simple cognitive tasks that involve minimal contributions from higher order functions (Fry & Hale, 2000). Strong correlational associations between processing speed and WM capacity have been well established, primarily in the literature on WM development (Kail & Park, 1994; Fry & Hale, 2000; Bayliss, Jarrold, Baddeley, Gunn and Leigh, 2005).

Slower processing speed is a well-established finding in children with ADHD. RTs of children with ADHD on a variety of tasks are characteristically slower and more variable than those of typically-developing children (Castellanos et al., 2005; Hervey et al., 2006; Klimkeit, Mattingly, Sheppard, Lee & Bradshaw, 2005; Holdnack, Moberg, Arnold, Gur & Gur, 1995; Katz, Brown, Roth & Beers, 2011; Shanahan et al., 2006). Effects of childhood ADHD on RT speed and variability are typically of moderate to large size (Karalunas et al., 2014), appear to be a heritable feature of the disorder (Andreou et al., 2007; Kuntsi et al., 2006) and are reduced by stimulant medication treatment (Kofler et al. 2013). Recent studies have largely supported the view that slow and variable RTs in ADHD are due to less efficient central processing speed (Huang-Pollock, Karalunas, Tam & Moore, 2012; Karalunas, Huang-Pollock and Nigg, 2012; Metin et al., 2013; Weigard & Huang-Pollock 2014) and that processing speed partially mediates the relationship between ADHD status and WM ability (Karalunas and Huang-Pollock, 2013).

Thus, preliminary evidence suggests that ADHD-related WM deficits could be partially explained by slower processing speed in this clinical population. Consideration of processing speed, as a basic and parsimoniously defined cognitive capacity, may thus be essential for the development of strong, mathematically-specified models of WM deficits in

ADHD. However, empirical support for this conjecture is limited by two features of prior work. The first is that the work linking processing speed to WM deficits in ADHD and other populations (Karalunas and Huang-Pollock, 2013) is correlational in nature. No prior study has experimentally evaluated whether slowing processing speed can directionally reduce WM recall in an ADHD population. The second is that it is unclear exactly *why* processing speed should limit WM recall in ADHD, as a mechanistic explanation for this directional relationship has not previously been proposed. In the current study, we seek to address these limitations by 1) experimentally manipulating processing speed in a complex-span WM paradigm to examine whether this manipulation impacts the WM recall of children with ADHD and their typically-developing peers and 2) determining whether a strong current model of WM capacity limits, the time-based resource-sharing (TBRS) model (Portrat, Camos & Barrouillet, 2009) can provide a mechanistic explanation for how slow processing speed may limit WM capacity in ADHD.

The time-based resource-sharing model

The TBRS model (Barrouillet, Bernardin & Camos, 2004; Portrat, Camos & Barrouillet, 2009) posits that items stored in WM decay unless they are frequently refreshed by a capacity-limited attentional focus (i.e., attentional “bottleneck”) which is prevented from refreshing memory items while it is taken up by the processing of concurrent tasks. Thus, this model predicts that “cognitive load”, mathematically defined as the proportion of time that the attentional bottleneck is occupied by processing concurrent tasks (relative to time it is available for refreshing WM items), should have a negative linear relationship with WM recall performance (Barrouillet & Camos, 2012). As cognitive load increases and the proportion of time available for refreshing decreases, more WM items are lost to decay. Within this framework, more efficient processing speed would decrease cognitive load by expediting the completion of concurrent tasks, which in turn would prevent decay.

Support for the TBRS model comes from studies in which parameters of “complex span” paradigms, a widely-used measure of WM and the workhorse of individual differences research (Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005; Daneman & Carpenter, 1980), are experimentally manipulated. In these tasks, the serial presentation of memory items is regularly interspersed with a secondary, attention-demanding “distractor” task. Whereas a “simple span” task only involves the presentation of memory items, such as words or numbers, and thus indexes short term storage ability, the secondary distractor task in complex span paradigms fulfills the construct's requirement that a WM measure involve concurrent storage and processing of information (Conway et al., 2005). For example, in the symmetry span task (Unsworth, Heitz, Schrock, & Engle, 2005; Unsworth, Redick, Heitz, Broadway, & Engle, 2009), participants are asked to correctly recall the location of spatial targets in the order they appear. In between the presentation of each to-be-remembered target, a geometric figure appears, and participants are asked to determine whether the figure is or is not symmetrical.

When the amount of time available in between the presentation of memory items is increased or the length of the secondary task itself is decreased, both of which necessarily decrease cognitive load, memory recall improves (Barrouillet, Bernardin & Camos, 2004;

Barrouillet & Camos, 2012). For example, Gaillard, Barrouillet, Jarrold, & Camos (2011) demonstrated that when 11-year old children are given a more difficult distractor task than 9-year-old children (adding 2 to a digit in each interval rather than adding 1), it equated the time 11 year olds spent completing the distractor to that of 9 year olds, thus equating cognitive load, and age-based differences in memory recall were greatly reduced. This supports the TBRS explanation that, because older children typically process distractor tasks in shorter amounts of time, they experience less cognitive load than younger children, leading to better WM recall than their younger counterparts.

The finding carries clear implications for the relationship between processing speed and WM recall in ADHD. If children with ADHD take longer to process a given concurrent task than their typically-developing peers, then they experience higher levels of cognitive load, which would explain their lower WM recall. The current study seeks to test this explanation by exploring whether an experimental effect of processing speed on WM performance in ADHD can be attributed to increases in cognitive load and whether group differences in cognitive load could explain baseline WM differences between children with ADHD and their typically developing peers.

Diffusion model-based measurement of speed and cognitive load

A methodological limitation of the broader literature on processing speed that the current study seeks to address is that most previous studies have indexed “speed” using mean RT. This measurement method contaminates the construct of interest, central cognitive processing efficiency, with other processes, such as motor response time, lower level perceptual processes, and strategy effects (e.g. emphasizing speed over accuracy when responding), all of which contribute to the mean RT (Ratcliff & McKoon, 2008). Thus, using mean RT can obscure the relationship between central cognitive processing efficiency and WM recall because of the influence of these extraneous factors on RT. However, the use of formal “evidence accumulation” models to operationalize speed provides a potential solution to this issue.

Evidence accumulation models are biologically-informed mathematical models that explain inter- and intra-individual latency and variability in RT, and have contributed to research on both the neural processes that underlie decision making (Smith & Ratcliff, 2004) and the role of processing efficiency in cognition (Schmiedek et al., 2007). Of these, the Ratcliff drift diffusion model (Ratcliff & McKoon, 2008) frames a two-choice decision as a single evidence accumulation process that drifts between two response boundaries (correct and incorrect) until a response is generated by contact with one (See Figure 1). The drift rate (v) is the average rate at which evidence for a decision is accumulated (e.g., whether a string of letters is a word or non-word, or whether a visual array contains a large or small number of stimuli). Drift rate varies with both experimental manipulations (more difficult tasks slow drift rate relative to easier tasks), and with individual differences in speed of processing (Ratcliff & McKoon, 2008; Ratcliff, Love, Thompson & Opfer, 2012). The parameter indexing the distance between the correct and error response boundaries, boundary separation (a), accounts for speed-accuracy tradeoffs; individuals who seek less confirmatory evidence before initiating a decision have narrower boundaries, allowing the

diffusion process to terminate at a boundary sooner, but leading to more terminations at the incorrect boundary from noise. The model also includes a parameter that indexes the time spent completing processes not associated with cognitive processing, non-decision time (T_{er}). These non-decisional processes are assumed to include the encoding of stimuli and motor preparation, but in theory, T_{er} accounts for all processes, cognitive, perceptual or otherwise, that occur during the RT but are not directly involved in the decision process itself (Ratcliff & McKoon, 2008).

Using the diffusion model to operationalize processing efficiency has several key advantages over the analysis of mean RT. First, factoring out components of RT not related to the processing of the decision (i.e., T_{er}) removes portions of RT related to extraneous non-decisional processes. Second, the a parameter controls for possible confounds related to individuals' speed-accuracy trade-off settings, which, if not considered, may bias results or weaken relationships between speed and other measures. In an empirical example, older adults have been found to display longer reaction times not because of slower drift rates, but because of conservative responding (wider boundaries) (Ratcliff, Thapar & McKoon, 2004). Indeed, these measurement advantages have already contributed to the body of work linking processing speed to WM ability; studies using evidence accumulation models to operationalize central cognitive processing speed have revealed strong, selective correlational relationships between the drift rate parameter and WM ability, both in adults (Ester, Ho, Brown & Serences, 2014; Schmiedek et al., 2007) and children with ADHD (Karalunas & Huang-Pollock, 2013).

Another strength of using the diffusion model to operationalize processing speed that is crucial to the current investigation is that the model also allows for the estimation of cognitive load within the same framework. The total time needed to complete the secondary distractor task in a WM span paradigm can be divided into: (a) time required to encode the stimulus (T_{er}), (b) time required to accumulate evidence to make a decision (decision time), and (c) time required to prepare a motor response (also T_{er}) (Figure 1). Previous empirical work has shown that the evidence accumulation process during even simple tasks competes for the same attentional resource needed for the maintenance of memory items (Ester, Ho, Brown & Serences, 2014). This suggests that, consistent with the “bottleneck” assumption of the TBRS model, memory items cannot be refreshed during the “decision time” ($RT - T_{er}$) component of RT. Thus, though most commonly conceptualized as indexing stimulus encoding and motor preparatory time, T_{er} should also represent the portion of the RT that is spent refreshing memory items, particularly in a situation in which no additional time is allotted for refreshing (i.e., when there are no explicit gaps left between stimuli; Figure 1). This assumption is not without precedent, as previous studies utilizing evidence accumulation models have similarly proposed and demonstrated that T_{er} can capture WM operations, including task rule retrieval (Shahar Teodorescu, Usher, Pereg & Meiran, 2014; Schmitz & Voss, 2012). In a complex span task where no time is explicitly left open for refreshing between stimuli (as in the current experiment), this assumption allows cognitive load in each condition and for each individual to be estimated by subtracting the proportion of the mean RT taken up by nondecision processes from 1 (cognitive load = $1 - T_{er}/RT$). These estimates of cognitive load can then, in turn, be used to test the predictions of the TBRS model.

Current study

In sum, previous correlational work has suggested that slow processing speed may play a significant role in ADHD-related WM deficits. The current study is designed to address prior limitations in the literature by directly testing the degree to which *experimentally manipulating* processing speed, as indexed by drift rate in the diffusion model, directionally influences WM ability in children with and without ADHD and by testing the TBRS model as a mechanistic explanation for *why* this directional relationship may occur. We use a variation of a well-validated complex span task in which we systematically increase or decrease the efficiency with which the secondary distractor task can be processed and leave no time explicitly open for refreshing in order to allow measurement of the effects of this manipulation on both processing speed and cognitive load in the same model-based framework. By utilizing an evidence accumulation model, these processes can be observed independently of speed-accuracy trade off effects and other components of RT. Specific hypotheses are as follows:

Hypothesis 1: In replication of previous correlational research on processing efficiency and WM (Ester et al., 2014; Schmiedek et al., 2007; Karalunas & Huang-Pollock 2013), individual differences in v will be related to individual differences in WM recall.

Hypothesis 2: A directional relationship between slowed speed and lower WM recall would be supported if the experimental reduction of processing speed (v) negatively impacts WM recall. It is predicted that this effect will be seen in children with ADHD as well their typically-developing peers, although children with ADHD are expected to show lower baseline levels of both v and WM recall.

Hypothesis 3: The TBRS model argues that slow processing speed limits WM recall by increasing the cognitive load of a task, in which cognitive load is defined by the proportion of time the attentional bottleneck is occupied by processing concurrent tasks (as opposed to refreshing memory items). If the TBRS model can be used to explain why slow speed may limit WM recall in ADHD, then: (a) experimentally reducing processing speed should be found to increase cognitive load (defined as $1 - Ter/RT$), (b) cognitive load should be greater among children with ADHD vs. non-ADHD controls because of their slower speed, and (c) a linear regression model in which the average cognitive load for each condition in each group determines the average WM recall should describe the data well, and children with ADHD should simply fall lower than controls on the same regression line.

Methods

Sample

Children ages 8 through 12 with ($N=71$, 46 male) and without ($N=27$, 10 male) ADHD, were community recruited as part of an ongoing study on attention and learning processes. Children identified as having ADHD were required to meet DSM-IV criteria for ADHD including age of onset, duration, cross situational severity, and impairment as determined by a parental report on the Diagnostic Interview Schedule for Children version IV (DISC-IV) (Shaffer, Fisher, & Lucas, 1997). At least one parent and one teacher report of behavior on

the Attention, Hyperactivity, or ADHD subscales of the Behavioral Assessment Scale for Children (BASC-2: Reynolds & Kamphaus, 2004) or the Conners' Rating Scales (Conners': Conners, 2001) were also required to exceed the 85th percentile (T-score>61). Children prescribed a psychostimulant medication were required to cease taking their medication 24-48 hours in advance of the day of testing (mean washout time = 80.4 hours, median = 58 hours). One child in the study (ADHD male) was excluded prior to data analysis because of noncompliance while completing the task.

Non-ADHD controls had never been diagnosed or treated for ADHD in the past. They did not meet criteria for ADHD on the DISC-IV and were below the 80th percentile (T-score = 58) on all of the above listed rating scales. To equate IQ levels between the ADHD and control groups so that group differences could not be attributed to differences in general intelligence, potential non-ADHD controls with an estimated IQ>115 were excluded from participating in the study. In addition, children in both groups with an estimated IQ<80 were also excluded.

The presence of common childhood disorders, such as anxiety, depression, oppositional defiant disorder, and conduct disorder was assessed using the DISC-IV, but the presence of these conditions was not exclusionary. The sample demographics (reflecting regional demographics) were as follows: 72.7% Caucasian/non-Hispanic, 8.1% Caucasian/Hispanic, 2% other Hispanic, 7.1% African American, 1% Asian, 6.1% mixed and 3% unknown/missing. The average annual income of families in the sample fell between \$60,000 and \$80,000.

Procedure

Participants completed the experimental paradigm and cognitive screening measures as part of a larger battery of cognitive tasks. Parents were compensated with a \$100 gift card, provided with verbal clinical feedback on relevant test results, and children were allowed to choose a small toy (<\$2) from a prize box. A 2-subtest short form (Vocabulary, Matrix Reasoning) of the Wechsler Intelligence Scale for Children—IV (WISC-IV: Wechsler, 2003) provided an estimated IQ. The correlation of the 2 subtest short form with the full 12-subtest battery is 0.87 (Sattler, 2008).

Experimental Tasks

Children completed a simple span task and a complex span task, described below. The simple and complex span tasks were modified from the symmetry span task obtained from Randall Engle and colleagues (Unsworth, Heitz, Schrock, & Engle, 2005).

Simple Span—Children viewed a 4x4 grid in which one square at a time (the target) randomly turned red for 1800ms. The number of to-be-remembered targets varied from two to nine, with three trials presented at each set size. Trials were presented in random order, and all trials were administered. Children were asked to recall the correct spatial location of the targets in the order in which they appeared, and made their responses using a computer mouse on a blank grid. The partial credit load scoring system was used, in which children

received 1 point for each target correctly recalled in the correct position (Conway et al., 2005).

Following the simple span task and prior to the complex span task, children then completed a 100-trial practice round consisting solely of numerosity decision trials (described below) to familiarize them with that portion of the task. Data are not reported for these practice trials.

Complex span—Here, the presentation of to-be-recalled spatial targets was interleaved with a secondary distractor task (numerosity discrimination; Ratcliff & McKoon, 2008; Ratcliff, Love, Thompson & Opfer, 2012) lasting 3000ms (see Figure 2). First, a to-be-remembered spatial target was presented for 2000ms. Then, to prevent active rehearsal (Conway et al., 2005), children were asked to respond with a right or left mouse click, whether a 10×10 array of asterisks appearing in an invisible grid within a square box had “a lot” (i.e. >50, left mouse click) or “a little bit” (i.e. <50, right mouse click) of candy (i.e. the numerosity judgment). They were to *accurately complete as many numerosity judgments as they could* in the space of 3000ms. Immediately after each response, a new numerosity judgment stimulus was presented until 3000ms had elapsed. At that point, the next to-be-remembered spatial target was immediately presented (even if the child had not yet responded to the last item) for 2000ms, and so forth. After all targets were presented, children were asked to recall the spatial location of the targets in the order they were presented on a blank grid with the computer mouse. The partial credit load scoring system (Conway et al., 2005) was used as the dependent variable of memory recall. No feedback was given on the numerosity decisions in this block because errors are more frequent for difficult decisions, and could become a confounding factor. The 3000ms cutoff procedure was employed to ensure that all secondary distractor task trials were of the same length, and specifically that trials with difficult decisions did not last longer than those with easy decisions which could have influence recall by increasing the amount of time items had to be remembered (Towse & Hitch, 1995).

After the initial instructions had been delivered, children completed three practice trials at a set size of two. They then completed the experimental trials where the number of to-be-remembered targets varied from two to seven, with four trials presented at each set size, for a total of 24 recall trials. Trials were presented in random order, and all trials were administered. Half of the trials at each set size were randomly selected to be interspersed with difficult numerosity decisions. This was the *slow drift rate condition*, in which stimuli contained between 41-45 or 56-60 asterisks. The other half of the trials were interspersed with easier numerosity decisions (stimuli contained between 31-35 or 66-70 asterisks) and represented the *fast drift rate condition*. Altering decision difficulty in this manner is a validated method of manipulating drift rate, and the number of asterisks in each drift rate condition was adopted from prior research in typically developing children (Ratcliff et al., 2012). Summary statistics for the number of trials completed by participants in each group and condition are available in Supplementary Table 2. On average, children with ADHD completed 194.28 easy and 194.29 difficult trials while controls completed 200.96 easy and 198.79 difficult trials. The lowest number of trials completed by any participant in any condition was 109 trials (control group, easy), which is above the minimum number of trials

recommended for using the Kolmogorov-Smirnov method to fit the diffusion model (Voss, Nagler & Lerche, 2013).

Diffusion Model Fitting

The response time and accuracy data of the numerosity decisions were fit to the diffusion model to obtain estimates of the v , a and Ter parameters for each participant using the Fast-dm modeling program (Voss & Voss, 2007) downloaded from the authors' website: <http://www.psychologie.uni-heidelberg.de/ae/meth/fast-dm>. Prior to fitting, RTs for fast guesses (<300ms) and outlier RTs (>3000) were excluded from analysis, following procedures used previously for successfully fitting a numerosity discrimination task to the diffusion model (e.g., Ratcliff et al., 2012). Fast-dm estimates diffusion model parameters for each condition by fitting a cumulative distribution function (CDF) of *observed* correct and errors trials to a CDF *predicted* by the best-fitting set of parameters. After initial parameter values are determined using the EZ-diffusion diffusion model program (Wagenmakers, Van Der Maas, & Grasman, 2007), fast-dm uses a simplex-downhill method to fit the predicted and observed distributions in three successive attempts with increasingly strict fit criteria until the best possible model-fit (as indexed by p-value) is achieved. As fast-dm was able to fit all participants' data to the model well (all $p > .16$), no participants were excluded from analysis due to poor model fit.

Planned Analyses

To replicate previous correlational research, v , a , and Ter , derived from the numerosity decision judgements embedded in the complex span task were entered as predictors in a multiple regression, and the total memory recall from the complex span block was entered as the dependent variable. The experimental design of the complex span task yielded a Diagnosis (2: ADHD, Control) \times Processing Speed (2: fast drift, slow drift) repeated-measures ANOVA, and variables of interest (v , cognitive load, WM recall) were analyzed with this design to test the main hypotheses. Finally, to test the second prediction of the TBRS model, a linear regression model was fit to the relationship between the mean WM recall and the mean cognitive load for each group in each condition. Results for all hypotheses did not change when they were corrected for multiple comparisons using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995).

Results

Preliminary Group Analyses

Supplementary Table 1a provides descriptive statistics. As would be expected, children with ADHD displayed more inattentive, $F(1,96)=525.94$, $p < .001$, and hyperactive/impulsive, $F(1,96)=107.37$, $p < .001$, symptoms than controls. There were no statistically significant group differences in FSIQ, $F(1,96)=2.09$, $p = .15$, or age, $F(1,96)=1.50$, $p = .22$.

Validation of Complex Span task

A span-type (2: Simple/Complex) \times Diagnosis (2: ADHD/Control) repeated measures ANOVA comparing the percentage of items recalled on the simple vs. complex span tasks confirmed that children recalled fewer items while concurrently performing the numerosity-

discrimination task, $F(1,96)=82.55$, $\eta^2=.46$, $p<.001$, thus validating the complex span task as a measure of WM. A smaller interaction effect was also discovered, $F(1,96)=5.48$, $\eta^2=.05$, $p=.02$, in which children with ADHD displayed greater performance decrements when the secondary task was added than typically-developing children. This interaction may reflect greater demands placed on control of attention or the secondary task's constraint on refreshing in the complex span condition, both of which may disproportionately affect individuals with ADHD.

Validation of the processing speed manipulation

Compared to the easy trials, the harder numerosity discrimination trials increased error rates, $F(1,96)=383.39$, $\eta^2=.80$, $p<.001$, increased standard deviations of RT, $F(1,96)=25.37$, $\eta^2=.21$, $p<.001$, decreased boundary separation, $F(1,96)=31.16$, $\eta^2=.25$, $p<.001$, and most importantly for validation of the manipulation, slowed drift rates, $F(1,96)=302.58$, $\eta^2=.76$, $p<.001$ (See Figure 3a and Supplementary Table 3). There was no significant main effect of difficulty on mean RT, $F(1,96)=3.12$, $\eta^2=.03$, $p=.08$, or Ter , $F(1,96)=3.56$, $\eta^2=.04$, $p=.062$. The lack of an effect on MRT was due to a decrease in boundary separation seen during difficult trials, indicating the presence of a strategic speed/accuracy tradeoff effect to compensate for slower drift for those trials. These analyses highlight the utility of the DDM to yield a more pure measure of processing speed through the isolation of confounding effects of speed/accuracy trade-offs on MRT.

Consistent with previous literature, children with ADHD were also less accurate, $F(1,96)=13.027$, $\eta^2=.12$, $p<.001$, had more variable RTs, $F(1,96)=6.89$, $\eta^2=.07$, $p=.01$, and displayed slower drift rates, $F(1,96)=14.36$, $\eta^2=.13$, $p<.001$, than their typically-developing peers. Diagnosis \times difficulty interactions were seen for both accuracy, $F(1,96)=9.40$, $\eta^2=.09$, $p=.003$ and drift, $F(1,96)=13.04$, $\eta^2=.12$, $p<.001$, in which group differences were larger in the easy vs. difficult conditions. There were no significant interactions for RT, SDRT, boundary separation, Ter (all $p > 0.18$, all $\eta^2 < 0.02$).

Hypothesis 1. Regression Analyses predicting WM capacity from diffusion model parameters. The diffusion model parameters derived from the numerosity decision trials were used to predict recall on the complex span task. Zero-order correlations as well the β weights are reported in Supplementary Table 1b. Diffusion model parameters v , a , and Ter , explained a significant portion of variance in WM recall in the complex span task, $F(3,97)=5.64$, $R^2=.15$, $p=.001$. However, consistent with previous literature, only drift rate predicted a significant proportion of unique variance in WM capacity, $\beta=.41$, $t(97)=4.06$, $p<.001$.

Hypothesis 2. Experimental effect of processing speed on WM. Supplementary Table 3 shows all main effects and interactions. As expected, children with ADHD had worse WM recall than controls, $F(1,96)=23.61$, $\eta^2=.20$, $p<.001$, recalling on average 25% of spatial targets compared to 43% for controls. Consistent with study hypotheses, when drift rate was slowed, WM capacity decreased, $F(1,96)=5.13$, $\eta^2=.05$, $p=.026$ (Figure 3b). No interaction between speed condition and diagnostic status was detected, $F(1,96)=.04$, $\eta^2<.001$, $p=.83$, suggesting that, for both groups, working memory capacity was similarly affected by changes in processing speed.

Hypothesis 3. *TBRS as an explanatory model for why slow speed may limit WM recall in ADHD.* Consistent with the TBRS model, experimentally slowing drift rate led to higher cognitive load ($1 - Ter/RT$), $F(1,96)=9.90$, $\eta^2=.09$, $p=.002$, and children with ADHD experienced higher cognitive load, $F(1,96)=5.14$, $\eta^2=.05$, $p=.026$, than controls did on the same task (Figure 3b). The group \times load interaction was not significant, $F(1,96)=.10$, $\eta^2<.001$, $p=.76$. Consistent with previous work (Barrouillet, Bernardin & Camos, 2004; Gaillard et al., 2011), a linear regression model fit to the mean cognitive loads and WM recall scores in both groups (Figure 4) had a negative slope and explained a large proportion of the variance (92%) in mean WM recall, $\beta = -3.09$, $F(1,2) = -4.91$, $R^2 = .92$, $p = .039$. Also consistent with the TBRS account, the data points from the ADHD group fell lower on this regression line than those from the control group.

Discussion

Past correlational research has found that individual differences in processing speed strongly predict WM ability in typically developing children (Bayliss, Jarrold, Baddeley, Gunn and Leigh, 2005; Fry & Hale, 2000; Kail & Park, 1994; Schmiedek et al., 2007), and that processing speed mediates WM deficits in ADHD (Karalunas & Huang-Pollock, 2013), suggesting that this basic and parsimoniously defined cognitive capacity may be a key construct in understanding WM deficits in this population. By combining a well-validated WM paradigm with evidence accumulation modeling methods, the current study is the first to experimentally test whether a directional relationship between speed and WM recall could exist in ADHD, and is the first to test a mechanistic explanation for why this relationship might occur. As has been reported elsewhere, we found that individuals with ADHD had smaller WM capacities (Kasper et al., 2012; Schoechlin & Engle, 2005; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005) and slower drift rates (e.g., Huang-Pollock et al., 2013; Karalunas et al., 2012; 2014; Metin et al., 2013; Weigard & Huang-Pollock 2014) than their same aged peers. And, in a direct replication of prior work (e.g. Ester et al., 2014; Karalunas & Huang-Pollock, 2013; Schmiedek et al., 2007), only the drift rate parameter was uniquely associated with WM ability.

Building upon that previous work, we further demonstrated that when drift rate was deliberately slowed via a within-subjects experimental manipulation, it led to decreased WM recall, both among children with ADHD and their typically-developing counterparts. This finding is the first indication, in a childhood ADHD population, that a directional relationship between speed and WM recall is plausible. It also serves to integrate the burgeoning literature on drift rate deficits in ADHD (Huang-Pollock, Karalunas, Tam & Moore, 2012; Karalunas, Huang-Pollock and Nigg, 2012; Metin et al., 2013; Weigard & Huang-Pollock 2014) with causal explanations for WM deficits in the disorder. Consistent with the TBRS model (Barrouillet, Bernardin & Camos, 2004; Portrat, Camos & Barrouillet, 2009), additional analyses found that reductions in speed directionally impacted WM recall by increasing cognitive load, that children with ADHD experienced higher cognitive load than controls in the same task, and that WM recall in the whole sample was a negative linear function of cognitive load. Thus, this set of findings not only demonstrates that such a directional relationship is plausible, but also provides a mechanistic explanation, via the TBRS model, for why the relationship occurs. Namely, that slow processing speed in ADHD

impacts WM recall by increasing the time that must be spent processing secondary tasks relative to the time available for preventing decay by refreshing memory items. This explanation may be used as a starting point for refining current models of WM and other EF deficits in the disorder to include consideration of processing speed, and the building of new models that can be specified formally.

What mechanisms could cause processing speed deficits in ADHD in the first place? Individual differences in processing speed have been linked to the strength of white matter connections between brain regions (Madden et al., 2012; Wozniak & Muetzel, 2011; Bava, Jacobus, Mahmood, Yang & Tapert, 2010; Karbasforoushan, Duffy, Blackford & Woodward, 2015; Turken et al., 2008). Abnormal structural connections have also been found in ADHD (Hamilton et al., 2008; Pavuluri et al., 2009; Silk, Vance, Rinehart, Bradshaw & Cunningham, 2009), supporting the intuitive conclusion that processing speed deficits and their effects on complex cognitive processes are caused by basic features of neural structures which support efficient transfer of information.

There are also several indications that specific attentional functions could underlie processing speed. More attentive pre-decision states, as indexed by ERP correlates, have been linked to increases in the rate of evidence accumulation (Kelly & O'Connell, 2013), the guidance of visual attention towards relevant stimuli increases drift rate (Ho, Brown, Abuyo, Ku & Serences, 2012; Smith, Ratcliff & Wolfgang, 2004), and drift rates of children with ADHD are particularly slow when performance-monitoring processes need to be engaged (Weigard, Huang-Pollock & Brown, 2016). Although these findings may suggest that processing speed is little more than another index of the broad “executive attention” construct, at least two strong, causal models have been proposed to explain the relationship between specific attentional mechanisms and processing speed in ADHD. Karalunas et al. (2014) recently posited that slow drift rates in ADHD may be explained by Aston-Jones and Cohen's (2005) “adaptive gain” model of the locus coeruleus norepinephrine (NE) system. This model proposes that phasic NE release, modulated by the prefrontal cortex and anterior cingulate, reduces the neural signal-to-noise ratio by providing a temporal attentional filter (Aston-Jones & Cohen, 2005). This filter selectively applies gain to neural processes that are relevant to the current task over those that are of less utility, and thus increases the efficiency of information processing, quantified by the authors as drift rate. A second theory, the neuroenergetic model of ADHD (Killeen, 2013; Killeen, Russell & Sergeant, 2013) suggests that tasks which require sustained attention/arousal over time or manipulations that increase cognitive demand (e.g. increasing event rate: Huang-Pollock, Ratcliff, McKoon, Shapiro, Weigard & Galloway-Long, 2016) deplete scarce neuronal metabolic resources, which leads to less efficient information processing. As the replenishment of neural energy is linked to the NE system, this theory posits that NE dysfunction limits the energy an individual with ADHD can bring to a given cognitive task, thus slowing drift rate (Killeen et al., 2013).

Both of these models involve the construct of attention but, crucially, both posit specific, causal and biologically-plausible mechanisms as explanations for slower processing speed in ADHD, and both are directly linked to formal diffusion models. Rather than referencing an attentional control homunculus, these theories highlight the power of basic processes, such as processing speed, and mathematically-specified models built to describe these processes,

such as the diffusion model, for allowing the role of attentional functions to be understood more mechanistically. More broadly, exploration of other basic processes that contribute to complex attention and cognitive control, such as signal detection, action selection, and others highlighted by Verbruggen et al. (2014), may eliminate the need to invoke higher-order EF constructs altogether. Given the substantial neuropsychological heterogeneity of children with ADHD (Fair, Bathula, Nikolas & Nigg, 2012), exploration of how these individual mechanisms impact WM and other putative EFs, whether directly or through their impact on processing speed, may lead to insights about how individuals with different underlying neurocognitive aberrations display similar impairments. The contribution of processing speed to WM highlighted in the current study is potentially just one example of how the study of such basic cognitive capacities, which are typically overlooked in the study of ADHD, may be crucially important for understanding broader dysfunction in the disorder.

Given that WM is relevant to a wide range of psychiatric disorders beyond ADHD, the current study has broader implications for the study of psychopathology in general. WM dysfunction has been linked to major depression (Gohier et al., 2009; Rose & Ebmeier, 2006), anxiety (Hayes et al., 2008; Vytal et al., 2016), schizophrenia (Abi-Dargham et al., 2002), and personality disorders (De Brito et al., 2013; Krause-Utz et al., 2012), among other clinical phenomena, and preliminary evidence suggests that WM impairment is a transdiagnostic risk factor for general psychopathology in children (Huang-Pollock et al., revise & resubmit). The ubiquity of WM dysfunction in psychiatric disorders, combined with the fact that TBRS and several other well-specified formal models of WM (e.g., Oberauer et al., 2012) have now been proposed by cognitive scientists, provides a prime opportunity for clinical psychologists to use computational models to develop theories of psychopathology that are well-integrated with current cognitive neuroscience.

One caveat to our interpretation is that drift rate for children with ADHD in the easy condition was faster than that of controls in the difficult condition, but ADHD recall in the easy condition was not better than controls in the difficult condition. One reason this may be is that we assumed no refreshing could occur in the period of time that the diffusion process was unfolding, and while it dominated the focus of attention. This assumption is consistent with a long-standing tenet of cognitive psychology, that conscious or controlled cognitive processes such as decision-making and WM maintenance, occur serially (e.g. Shiffrin & Schnieder, 1977). In the context of WM, the understanding is that the completion of the secondary distractor task removes all items in WM from the focus of attention (Portrat et al., 2011; Unsworth & Engle, 2007), which was supported in our own data demonstrating worse recall in the complex vs. simple span tasks. However, if some children (particularly controls) were able to refresh some memory items during the actual decision process in the difficult condition, then it might have resulted in greater recall despite slower drift rates. Conversely, if children with ADHD were not able to refresh memory items during the decision process of the easy trials, then it could have resulted in worse recall despite faster drift rates.

Thus, although drift rate influences cognitive load, and was a key predictor of WM, future studies would need to find ways to control, measure, or otherwise manipulate the degree to which parallel processing can occur during the task to better understand how these non-decisional components interact. Future investigations might also demonstrate that drift rate

manipulations in a variety of different distractor tasks, including random dot motion, lexical decision-making, or color discrimination, have similar modulating effects on WM performance (i.e., a conjunction analysis). Finally, future studies could employ parametric manipulations of speed using several difficulty conditions to create a continuum of drift rates, and demonstrate that the manipulation of drift rate changes WM recall to the same degree at each level.

Such a parametric manipulation would also address another major limitation of the current data as it relates to testing predictions of the TBRS model, which is the fact that there were only two levels of cognitive load in each group. If each group displays several different levels of cognitive load in the various parametric conditions, a model in which two separate regression lines are fit to data in each group could be compared against the single regression model tested in this study. This would provide a stronger test of the TBRS model because it would confirm that the relationship between cognitive load and capacity is linear in both groups, and determine whether the slope or intercept of the function differs between them.

Of note, children in both groups demonstrated a speed-accuracy trade off effect in the slow drift rate condition as evinced by the reduction of the boundary separation parameter (Ratcliff & McKoon, 2008). Because the experimental manipulation altered boundary separation as well as drift rate, it could be argued that changes in boundary separation were responsible for changes in WM recall. However, unlike processing speed, no theoretical rationale or previous empirical support for such a mechanism exists, and correlational analyses found that of the DDM parameters, only drift rate uniquely predicted WM recall, making this possibility unlikely. Instead, this decrease in boundary separation with increasing difficulty can be seen as a strategic choice by participants to allow more time for refreshing. That is, if the amount of time children have to refresh memory items is reduced when processing speed is reduced, an effective counter-strategy (conscious or not) would be to become less cautious to shorten the amount of time spent processing the secondary task. However, to further rule out the possibility that changes in boundary separation might be driving changes in WM recall, future experiments might consider using adaptive response deadlines (e.g., Rinkenauer et al., 2004) to individually identify the deadlines needed to equate boundary separation across speed conditions in a pre-assessment.

Relatedly, while it has been argued by some (e.g., Conway et al., 2005) that individuals do not “trade off” between processing and storage functions in complex span tasks (for example, accepting poor performance on the secondary task in favor of putting more effort into memory maintenance), our results suggest that the opposite is true, and that consideration of these strategic effects should be taken into consideration during scoring. At a minimum, it is recommended that future studies utilizing complex span tasks examine and report secondary task performance and its relationships with other variables in addition to those of recall scores.

Conclusion

The current findings have several major implications for the study of WM deficits and other higher-order constructs in ADHD. First, taken together with previous research (Barrouillet &

Camos, 2012; Fry & Hale, 2000; Karalunas & Huang-Pollock, 2013; Schmiedek et al., 2007), they underscore the need to consider processing speed in causal models of ADHD-related WM deficits. Second, they suggest that the TBRS model (Portrat et al., 2009; Barrouillet & Camos, 2012; Gaillard, Barrouillet, Jarrold, & Camos, 2011), a well-established theory with clear mathematical predictions, may provide a useful starting point for doing so. More broadly, the results demonstrate that the investigation of basic and parsimoniously defined cognitive capacities is not simply helpful for understanding dysfunction in higher-order processes, but may in fact be essential for building better-specified theoretical models of this dysfunction. Thus, this study encourages the scientific community studying ADHD to extend its focus beyond higher-order executive functions (Alderson et al., 2013; Barkley, 1997; Burgess et al., 2010; Castellanos & Tannock, 2002; Coghill et al., 2005; Rapport et al., 2008; Willcutt et al., 2005), and towards the exploration of low-level neurocognitive mechanisms in the disorder.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by National Institute of Mental Health Grant R01 MH084947 to Cynthia Huang-Pollock. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health. We also thank the parents, teachers, and children who participated, and the tireless research assistants who helped in the conduct of the study.

References

- Abi-Dargham A, Mawlawi O, Lombardo I, Gil R, Martinez D, Huang Y, et al. Gorman JM. Prefrontal dopamine D1 receptors and working memory in schizophrenia. *The Journal of Neuroscience*. 2002; 22(9):3708–3719. [PubMed: 11978847]
- Alderson RM, Hudec KL, Patros CH, Kasper LJ. Working memory deficits in adults with attention-deficit/hyperactivity disorder (ADHD): An examination of central executive and storage/rehearsal processes. *Journal of abnormal psychology*. 2013; 122(2):532. [PubMed: 23421528]
- Alloway TP, Gathercole SE, Holmes J, Place M, Elliott JG, Hilton K. The diagnostic utility of behavioral checklists in identifying children with ADHD and children with working memory deficits. *Child psychiatry and human development*. 2009; 40(3):353–366. [PubMed: 19280339]
- Alloway TP, Gathercole SE, Kirkwood H, Elliott J. The cognitive and behavioral characteristics of children with low working memory. *Child development*. 2009; 80(2):606–621. [PubMed: 19467014]
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th, text rev. Washington, DC: 2000.
- Andreou P, Neale BM, Chen WAI, Christiansen H, Gabriels I, Heise A, et al. Kuntsi J. Reaction time performance in ADHD: improvement under fast-incentive condition and familial effects. *Psychological medicine*. 2007; 37(12):1703–1715. [PubMed: 17537284]
- Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci*. 2005; 28:403–450. [PubMed: 16022602]
- Baddeley A. Working memory: theories, models, and controversies. *Annual review of psychology*. 2012; 63:1–29.
- Baddeley, AD., Hitch, GJ. Working memory. In: Bower, GH., editor. *The psychology of learning and motivation*. Vol. 8. New York: Academic Press; 1974. p. 47-89.

- Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological bulletin*. 1997; 121(1):65. [PubMed: 9000892]
- Barkley RA, Grodzinsky G, DuPaul GJ. Frontal lobe functions in attention deficit disorder with and without hyperactivity: A review and research report. *Journal of abnormal child psychology*. 1992; 20(2):163–188. [PubMed: 1593025]
- Barrouillet P, Bernardin S, Camos V. Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*. 2004; 133(1):83. [PubMed: 14979753]
- Barrouillet P, Camos V. As time goes by temporal constraints in working memory. *Current Directions in Psychological Science*. 2012; 21(6):413–419.
- Bava S, Jacobus J, Mahmood O, Yang TT, Tapert SF. Neurocognitive correlates of white matter quality in adolescent substance users. *Brain and cognition*. 2010; 72(3):347–354. [PubMed: 19932550]
- Bayliss DM, Jarrold C, Baddeley AD, Gunn DM, Leigh E. Mapping the developmental constraints on working memory span performance. *Developmental Psychology*. 2004; 41(4):579–597.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995:289–300.
- Burgess GC, Depue BE, Ruzic L, Willcutt EG, Du YP, Banich MT. Attentional control activation relates to working memory in attention-deficit/hyperactivity disorder. *Biological psychiatry*. 2010; 67(7):632–640. [PubMed: 20060961]
- Castellanos FX, Sonuga-Barke EJ, Scheres A, Di Martino A, Hyde C, Walters JR. Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological psychiatry*. 2005; 57(11):1416–1423. [PubMed: 15950016]
- Castellanos FX, Tannock R. Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Reviews Neuroscience*. 2002; 3(8):617–628. [PubMed: 12154363]
- Carr LA, Nigg JT, Henderson JM. Attentional versus motor inhibition in adults with attention-deficit/hyperactivity disorder. *Neuropsychology*. 2006; 20(4):430. [PubMed: 16846261]
- Coghill D, Nigg J, Rothenberger A, Sonuga-Barke E, Tannock R. Whither causal models in the neuroscience of ADHD? *Developmental Science*. 2005; 8(2):105–114. [PubMed: 15720368]
- Conners, CK. *Conners' Rating Scales—Revised Technical Manual*. NY: Multi-Health Systems Inc; 2001.
- Conway AR, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*. 2005; 12(5): 769–786. [PubMed: 16523997]
- Daneman M, Carpenter PA. Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*. 1980; 19(4):450–466.
- Daneman M, Merikle PM. Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*. 1996; 3(4):422–433. [PubMed: 24213976]
- De Brito SA, Viding E, Kumari V, Blackwood N, Hodgins S. Cool and hot executive function impairments in violent offenders with antisocial personality disorder with and without psychopathy. *PloS one*. 2013; 8(6):e65566. [PubMed: 23840340]
- Engle RW. Working memory capacity as executive attention. *Current directions in psychological science*. 2002; 11(1):19–23.
- Engle RW, Kane MJ, Tuholski SW. Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. *Models of working memory: Mechanisms of active maintenance and executive control*. 1999:102–134.
- Engle RW, Tuholski SW, Laughlin JE, Conway AR. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*. 1999; 128(3):309. [PubMed: 10513398]
- Ester EF, Ho TC, Brown SD, Serences JT. Variability in visual working memory ability limits the efficiency of perceptual decision making. *Journal of vision*. 2014; 14(4):2.
- Fair DA, Bathula D, Nikolas MA, Nigg JT. Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences*. 2012; 109(17):6769–6774.

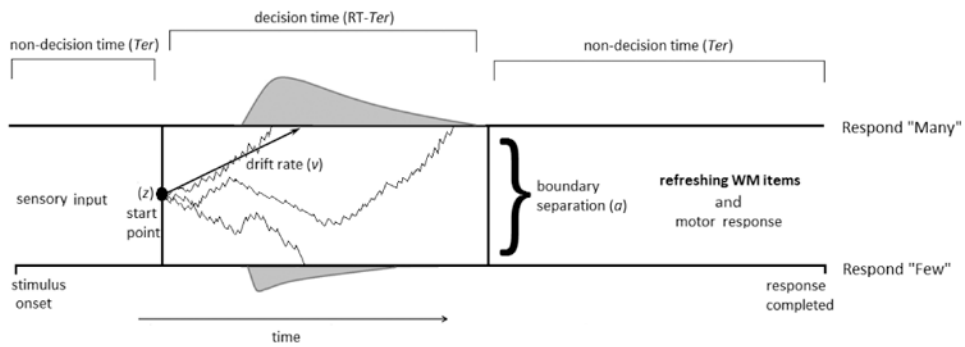
- Forstmann, BU., Wagenmakers, EJ. An Introduction to Model-Based Cognitive Neuroscience. Springer; New York: 2015. Model-Based Cognitive Neuroscience: A Conceptual Introduction; p. 139-156.
- Fry AF, Hale S. Relationships among processing speed, working memory, and fluid intelligence in children. *Biological psychology*. 2000; 54(1):1–34. [PubMed: 11035218]
- Gaillard V, Barrouillet P, Jarrold C, Camos V. Developmental differences in working memory: Where do they come from? *Journal of Experimental Child psychology*. 2011; 110(3):469–479. [PubMed: 21664622]
- Gathercole SE, Pickering SJ. Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology*. 2000; 70(2):177–194. [PubMed: 10900777]
- Geary DC, Hoard MK, Byrd-Craven J, Nugent L, Numtee C. Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*. 2007; 78(4):1343–1359. [PubMed: 17650142]
- Gohier B, Ferracci L, Surguladze SA, Lawrence E, El Hage W, Kefi MZ, et al. Le Gall D. Cognitive inhibition and working memory in unipolar depression. *Journal of affective disorders*. 2009; 116(1):100–105. [PubMed: 19042027]
- Goto Y, Hatakeyama K, Kitama T, Sato Y, Kanemura H, Aoyagi K, et al. Aihara M. Saccade eye movements as a quantitative measure of frontostriatal network in children with ADHD. *Brain and Development*. 2010; 32(5):347–355. [PubMed: 19505783]
- Gremillion ML, Martel MM. Semantic Language as a Mechanism Explaining the Association between ADHD Symptoms and Reading and Mathematics Underachievement. *Journal of abnormal child psychology*. 2012; 40(8):1339–1349. [PubMed: 22661106]
- Halin N, Marsh JE, Hellman A, Hellström I, Sörqvist P. A shield against distraction. *Journal of Applied Research in Memory and Cognition*. 2014; 3(1):31–36.
- Hamilton LS, Levitt JG, O'Neill J, Alger JR, Luders E, Phillips OR, et al. Narr KL. Reduced white matter integrity in attention-deficit hyperactivity disorder. *Neuroreport*. 2008; 19(17):1705. [PubMed: 18841089]
- Hayes S, Hirsch C, Mathews A. Restriction of working memory capacity during worry. *Journal of abnormal psychology*. 2008; 117(3):712. [PubMed: 18729625]
- Hervey AS, Epstein JN, Curry JF, Tonev S, Eugene Arnold L, Keith Conners C, et al. Hechtman L. Reaction time distribution analysis of neuropsychological performance in an ADHD sample. *Child Neuropsychology*. 2006; 12(2):125–140. [PubMed: 16754533]
- Ho TC, Brown S, Abuyo NA, Ku EHH, Serences JT. Perceptual consequences of feature-based attentional enhancement and suppression. *Journal of vision*. 2012; 12(8):15.
- Holdnack JA, Moberg PJ, Arnold SE, Gur RC, Gur RE. Speed of processing and verbal learning deficits in adults diagnosed with attention deficit disorder. *Cognitive and Behavioral Neurology*. 1995; 8(4):282–292.
- Huang-Pollock CL, Karalunas SL. Working memory demands impair skill acquisition in children with ADHD. *Journal of Abnormal Psychology*. 2010; 119(1):174. [PubMed: 20141254]
- Huang-Pollock CL, Karalunas SL, Tam H, Moore AN. Evaluating vigilance deficits in ADHD: A meta-analysis of CPT performance. *Journal of abnormal psychology*. 2012; 121(2):360. [PubMed: 22428793]
- Huang-Pollock C, Ratcliff R, McKoon G, Shapiro Z, Weigard A, Galloway-Long H. Using the Diffusion Model to Explain Cognitive Deficits in Attention Deficit Hyperactivity Disorder. *Journal of Abnormal Child Psychology*. 2016:1–12. [PubMed: 26687502]
- Huang-Pollock, Shapiro, Galloway-Long, Weigard. Are executive function deficits a transdiagnostic risk factor for psychopathology? *Journal of Abnormal Child Psychology*. revise and resubmit.
- Hughes RW, Hurlstone MJ, Marsh JE, Vachon F, Jones DM. Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*. 2013; 39(2):539. [PubMed: 22731996]

- Ikeda Y, Okuzumi H, Kokubun M. Stroop/reverse-Stroop interference in typical development and its relation to symptoms of ADHD. *Research in developmental disabilities*. 2013; 34(8):2391–2398. [PubMed: 23714715]
- Kail R, Park YS. Processing time, articulation time, and memory span. *Journal of experimental child psychology*. 1994; 57(2):281–291. [PubMed: 8169582]
- Kane MJ, Bleckley MK, Conway AR, Engle RW. A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*. 2001; 130(2):169. [PubMed: 11409097]
- Kane MJ, Engle RW. 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*. 2003; 132(1):47. [PubMed: 12656297]
- Kane MJ, Engle RW. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic bulletin & review*. 2002; 9(4):637–671. [PubMed: 12613671]
- Karalunas SL, Geurts HM, Konrad K, Bender S, Nigg JT. 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*. 2014; 55:685–710. [PubMed: 24628425]
- Karalunas SL, Huang-Pollock CL. Integrating impairments in reaction time and executive function using a diffusion model framework. *Journal of abnormal child psychology*. 2013:1–14. [PubMed: 22773360]
- Karalunas SL, Huang-Pollock CL, Nigg JT. Decomposing attention-deficit/hyperactivity disorder (ADHD)-related effects in response speed and variability. *Neuropsychology*. 2012; 26(6):684. [PubMed: 23106115]
- Karbasforoushan H, Duffy B, Blackford JU, Woodward ND. Processing speed impairment in schizophrenia is mediated by white matter integrity. *Psychological medicine*. 2015; 45(01):109–120. [PubMed: 25066842]
- Kasper LJ, Alderson RM, Hudec KL. Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): A meta-analytic review. *Clinical Psychology Review*. 2012
- Katz LJ, Brown FC, Roth RM, Beers SR. Processing speed and working memory performance in those with both ADHD and a reading disorder compared with those with ADHD alone. *Archives of clinical neuropsychology*. 2011; 26(5):425–433. [PubMed: 21613301]
- Kelly SP, O'Connell RG. Internal and external influences on the rate of sensory evidence accumulation in the human brain. *The Journal of Neuroscience*. 2013; 33(50):19434–19441. [PubMed: 24336710]
- Killeen PR. Absent without leave; a neuroenergetic theory of mind wandering. *Frontiers in psychology*. 2013; 4
- Killeen PR, Russell VA, Sergeant JA. A behavioral neuroenergetics theory of ADHD. *Neuroscience and Biobehavioral Reviews*. 2013; 37:625–657. [PubMed: 23454637]
- Klimkeit EI, Mattingley JB, Sheppard DM, Lee P, Bradshaw JL. Motor preparation, motor execution, attention, and executive functions in attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*. 2005; 11(2):153–173. [PubMed: 16036442]
- Kofler MJ, Alderson RM, Raiker JS, Bolden J, Sarver DE, Rapport MD. Working memory and intraindividual variability as neurocognitive indicators in ADHD: Examining competing model predictions. *Neuropsychology*. 2014; 28(3):459. [PubMed: 24588698]
- Kofler MJ, Rapport MD, Sarver DE, Raiker JS, Orban SA, Friedman LM, Kolomeyer EG. Reaction time variability in ADHD: a meta-analytic review of 319 studies. *Clinical psychology review*. 2013; 33(6):795–811. [PubMed: 23872284]
- Krause-Utz A, Oei NY, Niedtfeld I, Bohus M, Spinhoven P, Schmahl C, Elzinga BM. Influence of emotional distraction on working memory performance in borderline personality disorder. *Psychological Medicine*. 2012; 42(10):2181–2192. [PubMed: 22397907]
- Kuntsi J, Rogers H, Swinard G, Börger N, van der Meere JAAP, Rijdsdijk F, Asherson P. Reaction time, inhibition, working memory and 'delay aversion' performance: genetic influences and their interpretation. *Psychological Medicine*. 2006; 36(11):1613–1624. [PubMed: 16882357]

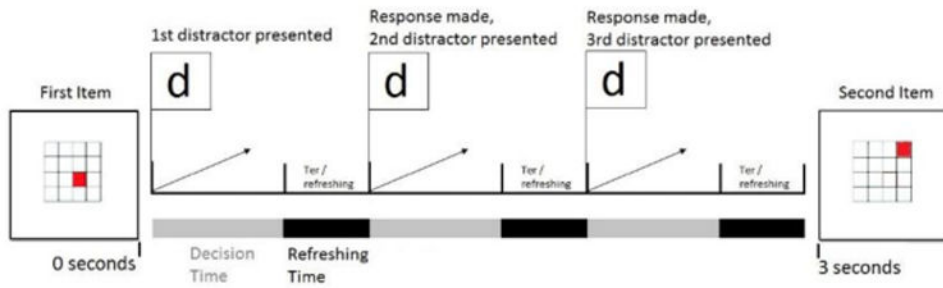
- Madden DJ, Bennett IJ, Burzynska A, Potter GG, Chen NK, Song AW. Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*. 2012; 1822(3):386–400. [PubMed: 21871957]
- Martinussen R, Hayden J, Hogg-Johnson S, Tannock R. A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2005; 44:377–384. [PubMed: 15782085]
- Meier M, Kane M. Working memory capacity and Stroop interference: Global versus local indices of executive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. May; 2013 39(3):748–759. 2013.
- Metin B, Roeyers H, Wiersema JR, van der Meere JJ, Thompson M, Sonuga-Barke E. ADHD performance reflects inefficient but not impulsive information processing: A diffusion model analysis. *Neuropsychology*. 2013; 27(2):193. [PubMed: 23527647]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*. 2000; 41(1):49–100. [PubMed: 10945922]
- Miller AC, Keenan JM, Betjemann RS, Willcutt EG, Pennington BF, Olson RK. Reading Comprehension in Children with ADHD: Cognitive Underpinnings of the Centrality Deficit. *Journal of abnormal child psychology*. 2013:1–11. [PubMed: 22773360]
- Miyake A, Friedman NP, Rettinger DA, Shah P, Hegarty M. How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*. 2001; 130:621–640. [PubMed: 11757872]
- Miyake, A., Shah, P. *Models of Working Memory: Mechanisms of Active Maintenance and Control*. New York: Cambridge University Press; 1999.
- Monsell S, Driver J. 1 Banishing the Control Homunculus. *Control of cognitive processes*. 2000:3.
- Nigg JT, Butler KM, Huang-Pollock CL, Henderson JM. Inhibitory processes in adults with persistent childhood onset ADHD. *Journal of consulting and clinical psychology*. 2002; 70(1):153. [PubMed: 11860041]
- Nyroos M, Wiklund-Hörnqvist C. The association between working memory and educational attainment as measured in different mathematical subtopics in the Swedish national assessment: primary education. *Educational Psychology*. 2012; 32(2):239–256.
- Oberauer K, Lewandowsky S, Farrell S, Jarrold C, Greaves M. Modeling working memory: An interference model of complex span. *Psychonomic bulletin & review*. 2012; 19(5):779–819. [PubMed: 22715024]
- Pavuluri MN, Yang S, Kamineni K, Passarotti AM, Srinivasan G, Harral EM, et al. Zhou XJ. Diffusion tensor imaging study of white matter fiber tracts in pediatric bipolar disorder and attention-deficit/hyperactivity disorder. *Biological psychiatry*. 2009; 65(7):586–593. [PubMed: 19027102]
- Portrat S, Camos V, Barrouillet P. Working memory in children: A time-constrained functioning similar to adults. *Journal of experimental child psychology*. 2009; 102(3):368–374. [PubMed: 18632113]
- Rapport MD, Alderson RM, Kofler MJ, Sarver DE, Bolden J, Sims V. Working memory deficits in boys with attention-deficit/hyperactivity disorder (ADHD): the contribution of central executive and subsystem processes. *Journal of Abnormal Child Psychology*. 2008; 36(6):825–837. [PubMed: 18317920]
- Ratcliff R, Love J, Thompson CA, Opfer JE. Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development*. 2012; 83(1):367–381. [PubMed: 22188547]
- Ratcliff R, McKoon G. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural computation*. 2008; 20(4):873–922. [PubMed: 18085991]
- Ratcliff R, Thapar A, McKoon G. A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*. 2004; 50(4):408–424.
- Reynolds, C., Kamphaus, R. *Behavior Assessment for Children, (BASC-2)*. Circle Pines, MN: American Guidance Service; 2004.

- Rinkenauer G, Osman A, Ulrich R, Müller-Gethmann H, Mattes S. On the locus of speed-accuracy trade-off in reaction time: inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General*. 2004; 133(2):261. [PubMed: 15149253]
- Rose EJ, Ebmeier KP. Pattern of impaired working memory during major depression. *Journal of affective disorders*. 2006; 90(2):149–161. [PubMed: 16364451]
- Sattler, J. *Resource Guide to Accompany Assessment of Children: Cognitive Foundations*. 5th. San Diego: Jerome Sattler Publisher, Inc; 2008.
- Schall JD. Neural correlates of decision processes: neural and mental chronometry. *Current Opinion in Neurobiology*. 2003; 13(2):182–186. [PubMed: 12744971]
- Scheiter K, Gerjets P, Heise E. Distraction during learning with hypermedia: difficult tasks help to keep task goals on track. *Frontiers in psychology*. 2014; 5
- Schmiedek F, Oberauer K, Wilhelm O, Suss HM, Wittmann WW. Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology General*. 2007; 136(3):414. [PubMed: 17696691]
- Schmitz F, Voss A. Decomposing task-switching costs with the diffusion model. *Journal of experimental psychology: human perception and performance*. 2012; 38(1):222. [PubMed: 22060144]
- Schoechlin C, Engel RR. Neuropsychological performance in adult attention-deficit hyperactivity disorder: meta-analysis of empirical data. *Archives of Clinical Neuropsychology*. 2005; 20(6): 727–744. [PubMed: 15953706]
- Sergeant J. The cognitive-energetic model: an empirical approach to attention-deficit hyperactivity disorder. *Neuroscience & Biobehavioral Reviews*. 2000; 24(1):7–12. [PubMed: 10654654]
- Sergeant JA. Modeling attention-deficit/hyperactivity disorder: a critical appraisal of the cognitive-energetic model. *Biological psychiatry*. 2005; 57(11):1248–1255. [PubMed: 15949995]
- Shaffer, D., Fisher, P., Lucas, R. *NIMH Diagnostic Interview Schedule for Children-IV*. New York: Ruane Center for Early Diagnosis, Division of Child Psychiatry, Columbia University; 1997.
- Shahar N, Teodorescu AR, Usher M, Pereg M, Meiran N. Selective influence of working memory load on exceptionally slow reaction times. *Journal of Experimental Psychology: General*. 2014; 143(5):1837. [PubMed: 25000446]
- Shanahan MA, Pennington BF, Yerys BE, Scott A, Boada R, Willcutt EG, et al. DeFries JC. Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of abnormal child psychology*. 2006; 34(5):584–601.
- Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*. 1977; 84(2):127.
- Silk TJ, Vance A, Rinehart N, Bradshaw JL, Cunnington R. White-matter abnormalities in attention deficit hyperactivity disorder: A diffusion tensor imaging study. *Human brain mapping*. 2009; 30(9):2757–2765. [PubMed: 19107752]
- Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. *Trends in neurosciences*. 2004; 27(3):161–168. [PubMed: 15036882]
- Smith PL, Ratcliff R, Wolfgang BJ. Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision research*. 2004; 44(12): 1297–1320. [PubMed: 15066392]
- Towse JN, Hitch GJ. Is there a relationship between task demand and storage space in tests of working memory capacity? *The Quarterly Journal of Experimental Psychology*. 1995; 48(1):108–124. [PubMed: 7754077]
- Turken U, Whitfield-Gabrieli S, Bammer R, Baldo JV, Dronkers NF, Gabrieli JD. Cognitive processing speed and the structure of white matter pathways: convergent evidence from normal variation and lesion studies. *Neuroimage*. 2008; 42(2):1032–1044. [PubMed: 18602840]
- Unsworth N, Engle RW. The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological review*. 2007; 114(1):104. [PubMed: 17227183]
- Unsworth N, Heitz RP, Schrock JC, Engle RW. An automated version of the operation span task. *Behavior research methods*. 2005; 37(3):498–505. [PubMed: 16405146]

- Unsworth N, Redick TS, Heitz RP, Broadway JM, Engle RW. Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*. 2009; 17(6):635–654. [PubMed: 19536691]
- Vandierendonck A. A working memory system with distributed executive control. *Perspectives on Psychological Science*. 2016; 11(1):74–100. [PubMed: 26817727]
- van Ravenzwaaij D, Brown S, Wagenmakers EJ. An integrated perspective on the relation between response speed and intelligence. *Cognition*. 2011; 119(3):381–393. [PubMed: 21420077]
- Verbruggen F, McLaren IP, Chambers CD. Banishing the control homunculi in studies of action control and behavior change. *Perspectives on Psychological Science*. 2014; 9(5):497–524. [PubMed: 25419227]
- Voss A, Nagler M, Lerche V. Diffusion models in experimental psychology: a practical introduction. *Experimental psychology*. 2012; 60(6):385–402.
- Voss A, Voss J. Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*. 2007; 39(4):767–775. [PubMed: 18183889]
- Vytal KE, Arkin NE, Overstreet C, Lieberman L, Grillon C. Induced-anxiety differentially disrupts working memory in generalized anxiety disorder. *BMC psychiatry*. 2016; 16(1):1. [PubMed: 26739960]
- Wagenmakers EJ, Van Der Maas HL, Grasman RP. An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*. 2007; 14(1):3–22. [PubMed: 17546727]
- Wechsler, D. Wechsler Intelligence Scale for Children, 4th Ed (WISC-IV) Technical and Interpretive Manual. San Antonio: Harcourt Brace; 2003.
- Weigard A, Huang-Pollock C. A diffusion modeling approach to understanding contextual cueing effects in children with ADHD. *Journal of Child Psychology and Psychiatry*. 2014
- Weigard A, Huang-Pollock C, Brown S. Evaluating the Consequences of Impaired Monitoring of Learned Behavior in Attention-Deficit/Hyperactivity Disorder Using a Bayesian Hierarchical Model of Choice Response Time. *Neuropsychology*. 2016
- Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological psychiatry*. 2005; 57(11):1336–1346. [PubMed: 15950006]
- Wozniak JR, Muetzel RL. What does diffusion tensor imaging reveal about the brain and cognition in fetal alcohol spectrum disorders? *Neuropsychology review*. 2011; 21(2):133–147. [PubMed: 21347880]



Easy distractor (d)



Hard distractor (D)

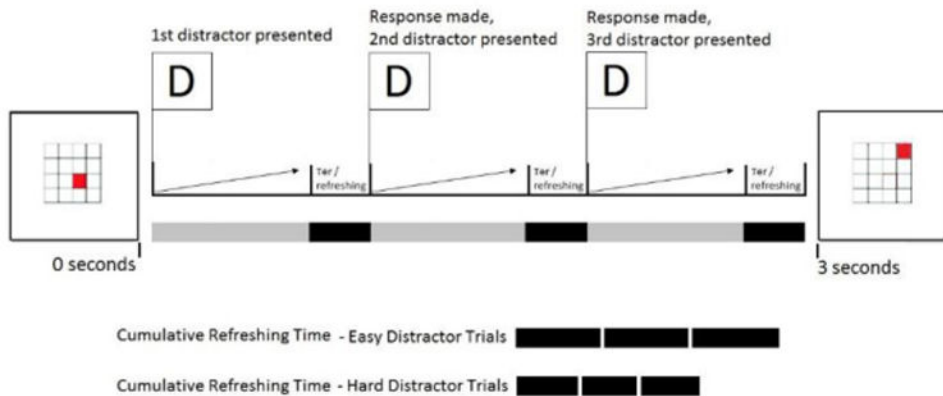


Figure 1.
a. The diffusion decision process (Ratcliff & McKoon, 2008) with memory refreshing processes included. In this example, participants must decide whether an array contains “many” (>50) or “few” (<50) asterisks. After the stimulus is presented, participants complete any necessary non-decision processes (sensory input, etc.), and then proceed to gather evidence for each alternative. Responses that terminate on the upper boundary register a response of “many” while those that terminate on the lower boundary register a response of “few”. Once the decision is made, additional time is taken to conduct the motor response and to refresh memory items.

b. The effect of slow drift on cognitive load and refreshing time. Children spend a portion of time processing the distractor (i.e. deciding many or few; “decision time,” gray bar) and spend a separate portion of time refreshing memory items (“refreshing time”; black bar). Hard distractor trials (D) slow drift relative to easy trials (d), increasing cognitive load, or the proportion of time spent processing the decision. When cognitive load increases and refreshing time is shortened (bottom).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

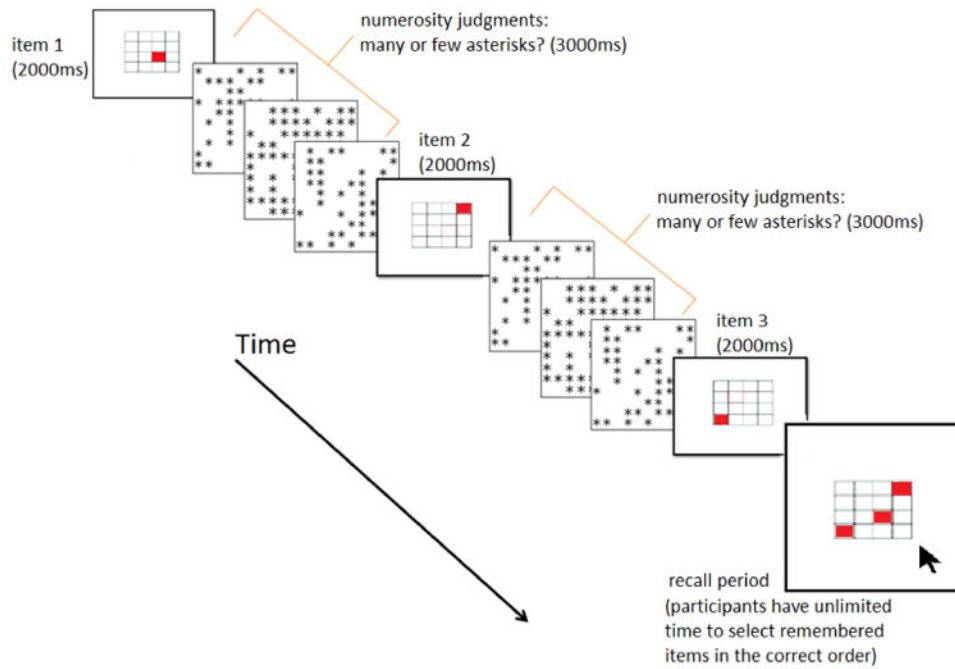
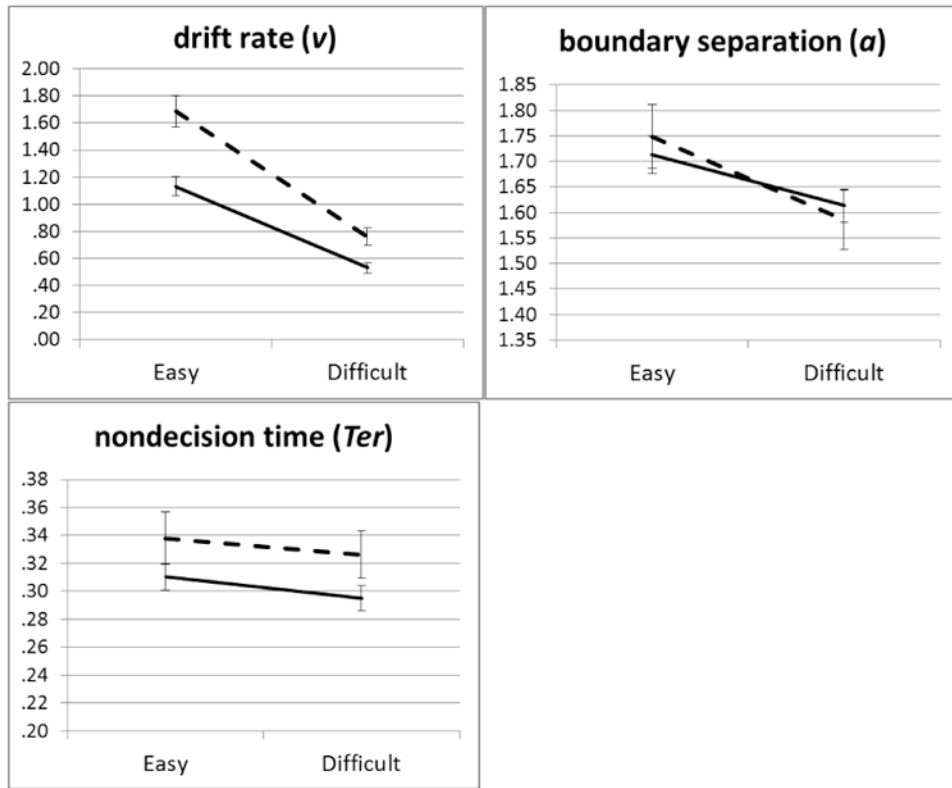


Figure 2.

Numerosity span task. Children were serially presented to-be-remembered spaces on a grid interspersed with numerosity decisions. Children completed as many numerosity decisions as quickly and accurately as they could in the time provided (3000ms). After presentation of the final to-be-remembered item, a blank grid appeared. Children used the mouse to indicate the order in which the targets appeared, and were given unlimited time to do so.



Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

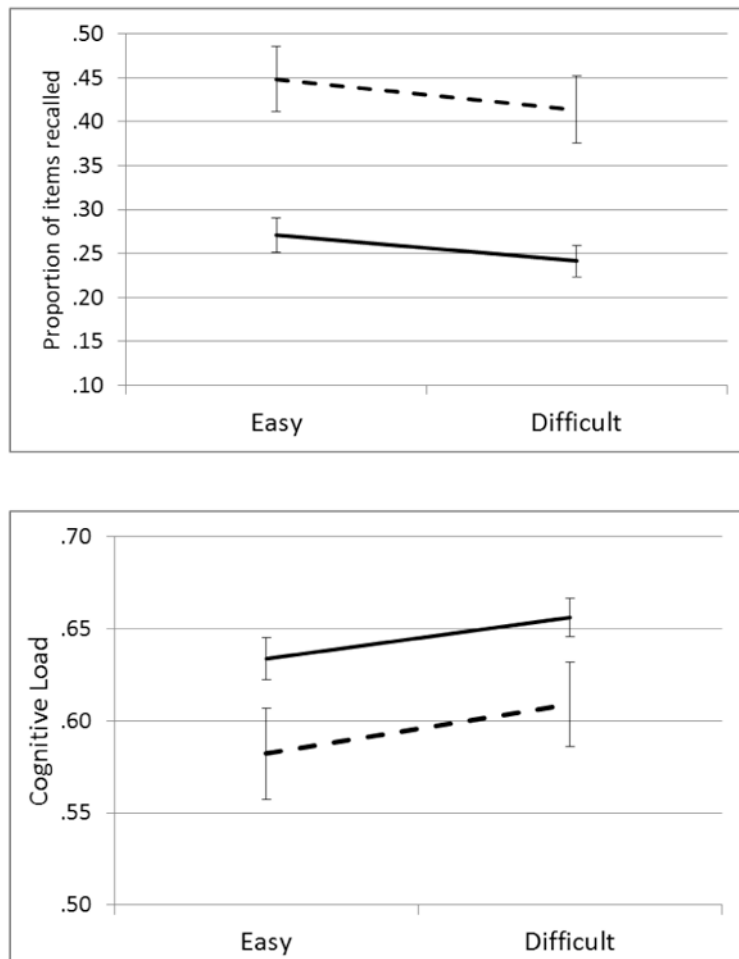


Figure 3.

a. Effects of secondary task difficulty (easy vs difficult) on diffusion model parameters during the complex span task by ADHD status (error bars reflect standard error of the mean). Control = dashed line, ADHD = solid line.

b. Effects of difficulty on WM performance and cognitive load ($1 - T_{er}/RT$) for refreshing (error bars reflect standard error of the mean); Control = dashed line, ADHD = solid line

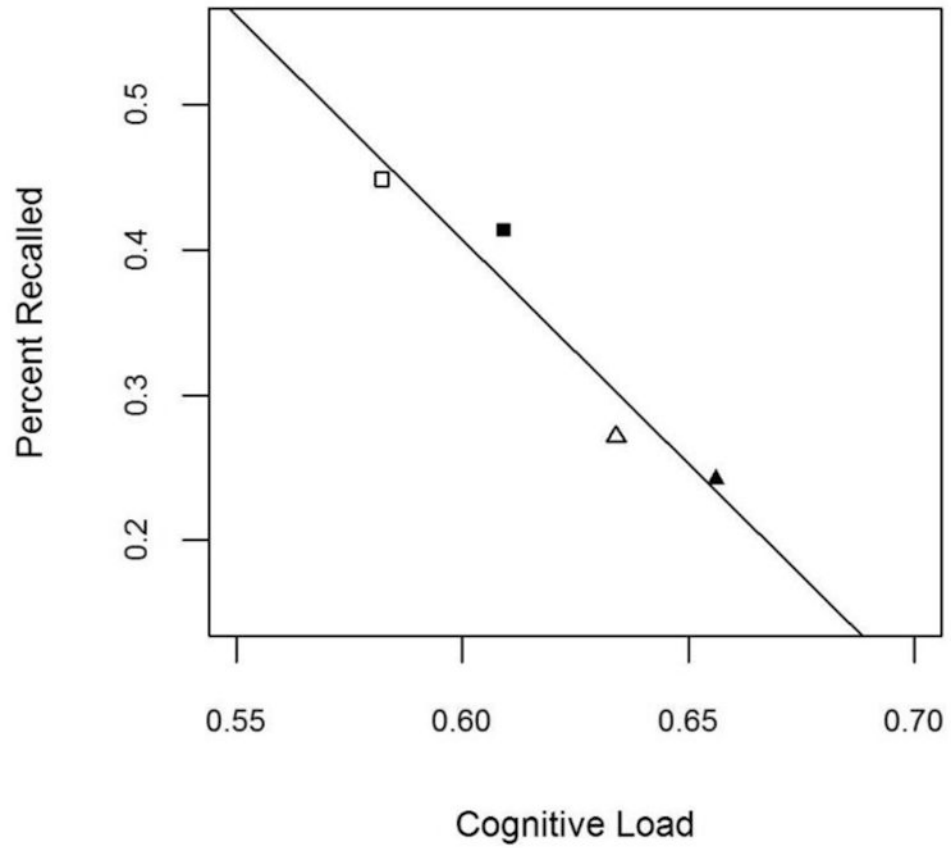


Figure 4. Mean memory recall in both groups and conditions as a negative linear function of cognitive load. Squares = control, Triangles = ADHD; Filled = slow drift, Empty = fast drift