Routledge
Taylor & Francis Group

Journal of Clinical and Experimental Neuropsychology

ISSN: 1380-3395 (Print) 1744-411X (Online) Journal homepage: https://www.tandfonline.com/loi/ncen20

Parent ratings of working memory are uniquely related to performance-based measures of secondary memory but not primary memory

Kathryn J. Ralph, Bradley S. Gibson & Dawn M. Gondoli

To cite this article: Kathryn J. Ralph, Bradley S. Gibson & Dawn M. Gondoli (2018) Parent ratings of working memory are uniquely related to performance-based measures of secondary memory but not primary memory, Journal of Clinical and Experimental Neuropsychology, 40:8, 841-851, DOI: 10.1080/13803395.2018.1441378

To link to this article: https://doi.org/10.1080/13803395.2018.1441378

Published online: 06 Mar 2018.

JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
2018, VOL. 40, NO. 8, 841–851
https://doi.org/10.1080/13803395.2018.1441378

Parent ratings of working memory are uniquely related to performance-based measures of secondary memory but not primary memory

Kathryn J. Ralph, Bradley S. Gibson and Dawn M. Gondoli

Department of Psychology, University of Notre Dame, Notre Dame, IN, USA

ABSTRACT

Introduction: Existing evidence suggests that performance- and rating-based measures of working memory (WM) correlate poorly. Although some researchers have interpreted this evidence as suggesting that these measures may be assessing distinct cognitive constructs, another possibility is that rating-based measures are related to some but not all theoretically motivated performance-based measures. The current study distinguished between performance-based measures of primary memory (PM) and secondary memory (SM), and examined the relation between each of these components of WM and parent-ratings on the WM subscale of the Behavior Rating Inventory of Executive Function (BRIEF–WM). Because SM and BRIEF–WM scores have both been associated with group differences in attention-deficit/hyperactivity disorder (ADHD), it was hypothesized that SM scores would be uniquely related to parent-rated BRIEF–WM scores.

Method: Participants were a sample of 77 adolescents with and without an ADHD diagnosis, aged 11 to 15 years, from a midwestern school district. Participant scores on verbal and spatial immediate free recall tasks were used to estimate both PM and SM capacities. Partial correlation analyses were used to evaluate the extent to which estimates of PM and SM were uniquely related parent-rated BRIEF–WM scores.

Results: Both verbal and spatial SM scores were significantly related to parent-rated BRIEF–WM scores, when corresponding PM scores were controlled. Higher verbal and spatial SM scores were associated with less frequent parent-report of WM-related failures in their child's everyday life. However, neither verbal nor spatial PM scores significantly related to parent-rated BRIEF–WM scores, when corresponding SM scores were controlled.

Conclusion: The current study suggested that previously observed low correlations between performance- and rating-based measures of WM may result from use of performance-based WM measures that do not capture the unique contributions of PM and SM components of WM.

ARTICLE HISTORY
Received 2 October 2017
Accepted 9 February 2018

KEYWORDS
Assessment; Behavior Rating Inventory of Executive Function; primary memory; secondary memory; working memory

Working memory (WM) is an integral construct in theories of cognition. As a system responsible for the temporary storage and retrieval of information during complex activities (Kane & Engle, 2002; Unsworth & Engle, 2007a), WM is important because it enables adaptive goal-directed behavior and strongly relates to higher order cognitive abilities such as fluid reasoning (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004). Consequently, researchers in both applied and basic settings have sought to devise reliable and valid measures of WM capacity.

Generally, WM capacity has been assessed using both "performance-based" and "rating-based" measures. On one hand, performance-based measures of an individual's WM capacity have been developed by basic researchers to test models of WM capacity; these are tasks completed by the individual under highly controlled and standardized laboratory conditions. For instance, immediate free recall (IFR) tasks are one type of performance-based measure that have been widely used to assess individual differences in WM and that have formed the basis for the dual-component model of WM (Unsworth & Engle, 2007a).

The dual-component model holds that WM is composed of primary memory (PM) and secondary memory (SM) components. PM refers to the active maintenance of a limited amount of goal-relevant information over the short term, specifically in the face of internal and external distraction or interference. However, the capacity of PM can be overwhelmed due to limitations of storage and/or attentional resources and, thus, lead to a disruption in goal-directed behavior. Nevertheless, goal-directed behavior may be

CONTACT Kathryn J. Ralph kralph@nd.edu Department of Psychology, University of Notre Dame, 106 Haggar Hall, Notre Dame, IN 46556, USA

resumed when the information lost from PM is recovered from a more enduring memory store known as episodic long-term memory or secondary memory (SM) using cue-dependent retrieval mechanisms (Mogle, Lovett, Stawski, & Sliwinski, 2008; Unsworth, 2007, 2009; Unsworth & Engle, 2007a, 2007b; Unsworth & Spillers, 2010). Importantly, estimates of both PM and SM using performance-based measures such as IFR tasks have been shown to reliably distinguish between high- and low-WM-capacity individuals (Unsworth & Engle, 2007a).

On the other hand, rating-based measures of an individual's WM capacity have been developed by applied researchers to assess everyday problems that arise from weaknesses in WM capacity (Rabin, Paolillo, & Barr, 2016); these are questionnaires with items describing various behaviors that are rated by an informant, such as a parent or teacher, who is well acquainted with the individual in question. For example, using factor analytic techniques, scales on the Behavior Rating Inventory of Executive Function (BRIEF) were derived by grouping together behavioral descriptors thought to be manifestations of underlying cognitive abilities (Gioia, Isquith, Guy, & Kenworthy, 2000). The BRIEF is composed of two major index scales called the Behavioral Regulation Index (BRI) and the Metacognition Index (MI), and a composite of these two scales called the Global Executive Composite (GEC). Each of the two major index scales is further composed of subscales. The BRI is composed of the Inhibit (IN), Shift (SH), and Emotional Control (EC) scales, and the MI is composed of the Initiate (IT), Working Memory (WM), Plan/Organize (PO), Organization of Materials (OM), and Monitor Scales (MO). Of critical interest in the present article is the BRIEF–WM subscale, which was designed to assess how well participants hold information in mind while completing goal-directed activities in everyday life (Gioia, Isquith, Retzlaff, & Epsy, 2002; see Faridi et al., 2015, for example items).

Although the BRIEF–WM scale is an ecologically reliable and valid measure for distinguishing children and adolescents with and without WM problems, this scale correlates poorly (and often nonsignificantly) with some performance-based measures (Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Mahone, Martin, Kates, Hay, & Horská, 2009; Mahone et al., 2002; McAuley, Chen, Goos, Schachar, & Crosbie, 2010; Miranda, Colomer, Mercader, Fernández, & Presentación, 2015; Toplak, Bucciarelli, Jain, & Tannock, 2009; Toplak, West, & Stanovich, 2013; Vriezen & Pigott, 2002). This finding has led some researchers to assert that performance- and rating-based measures provide different types of information about cognitive and behavioral functioning, assessing distinct processes or levels of a construct (Barkley & Fischer, 2011; Toplak et al., 2013). However, these low correlations may be driven by use of performance-based measures such as digit span or spatial span that do not adequately measure the SM component of WM (at least when the span lengths encountered by an individual are determined by the individual's ability, see Unsworth & Engle, 2006). Indeed, there is indirect evidence, obtained from studies of attention-deficit/hyperactivity disorder (ADHD) suggesting that BRIEF–WM scores may be more highly related to SM scores than to PM scores.

For instance, Gibson and colleagues (Gibson, Gondoli, Flies, Dobrzenski, & Unsworth, 2010; Gibson, Gondoli, Ralph, & Sztybel, 2018) used verbal and spatial IFR tasks and showed that individuals with confirmed diagnoses of ADHD were more deficient in SM abilities than PM abilities across both task modalities relative to individuals without ADHD. Furthermore, individuals with ADHD also typically have higher BRIEF–WM scores, indicating more WM problems, than individuals without ADHD (Mahone et al., 2002; Toplak et al., 2009). Hence, if ADHD scores are highly correlated with BRIEF–WM scores, then this pattern of findings can be interpreted to suggest that BRIEF–WM scores may be more highly related to SM scores than to PM scores.

Although the findings reported by Gibson and his colleagues (2010, 2018) are suggestive, there has been no direct measure of the relation between PM, SM, and BRIEF–WM scores. Accordingly, the present study was designed to directly test the hypothesis that the correlation between BRIEF–WM scores and SM scores would be more robust than the correlation between BRIEF–WM scores and PM scores. Estimates of the PM and SM components were measured by both verbal and spatial IFR tasks in a sample of adolescents with and without ADHD. Of critical importance, because estimates of PM and SM are themselves positively correlated (at least when recall initiation strategies are controlled, see Gibson & Gondoli, 2017), partial correlation was used to assess the extent to which the correlation between BRIEF–WM and SM would remain significant while controlling for PM in each task modality, and vice versa.

In summary, the present study sought to examine the extent to which performance-based and rating-based measures may target overlapping mechanisms once finer grained performance-based measures capable of measuring both the PM and SM components of WM are used.

Method

Participants

Participants were the same sample as that previously reported in Gibson et al. (2018). In that study, initial recruitment letters were mailed to adolescents from three middle schools (Grades 6–8) in a midwestern school district (approximately 2500 households). The letter sent to families described the researchers' interest in studying WM in adolescents both with and without a diagnosis of ADHD. The inclusion of those with and without ADHD was appropriate in this study to ensure that the range of scores was not restricted on the performance- and rating-based measures.

One hundred and nine families responded to the letter and were subsequently scheduled for a 2-hour session at a laboratory located in the Psychology Department at the University of Notre Dame. At this session, adolescents were assessed with the Wechsler Abbreviated Scale of Intelligence (WASI; The Psychological Corporation, 1999), the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), and a spatial and verbal IFR task. Parents completed the National Institute of Mental Health Computerized Diagnostic Interview Schedule for Children, Version 4 (NIMH C-DISC–IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) to confirm or deny a diagnosis of ADHD. Parents also completed a questionnaire regarding the type, time, and dosage of any medications that their child was taking, as well as the BRIEF and DuPaul ADHD Rating Scale (DuPaul, Power, Anastopoulos, & Reid, 1998). Parents and adolescents provided informed consent and assent, respectively, and in accordance with the study protocol approved by the institutional review board at the University of Notre Dame.

At recruitment, 56 participants reported having an ADHD diagnosis, and 52 participants reported having no ADHD diagnosis. One participant was excluded for having a full-scale IQ below 70. Participants were also excluded if their diagnostic status at recruitment did not match their diagnosis based on the C-DISC–IV. Following parent interviews, eight participants that had originally identified themselves as having an ADHD diagnosis but did not meet diagnostic criteria for ADHD were excluded. All participants that identified as being ADHD free at recruitment did not meet criteria for ADHD diagnosis (n = 52). Therefore, 100 participants (n = 48 with ADHD, n = 52 without ADHD) were assessed with IFR tasks.

Of the 48 participants in the ADHD group, 32 were currently taking stimulant medication (dexmethylphenidate, lisdexamfetamine, methylphenidate, and/or mixed amphetamine salts), five were taking nonstimulant medication (atomoxetine, guanfacine), one was being treated with non-ADHD medication (bupropion), and 10 were not being medicated. All ADHD medication was withheld at least 24 hours prior to testing. No participants without an ADHD diagnosis were currently medicated.

Parent rating-based measure

Primary caregivers completed the BRIEF, an 86-item behavioral rating questionnaire designed to assess deficits in child and adolescent executive functioning in their natural, everyday environment (Gioia et al., 2000; Toplak et al., 2013). For each item on the BRIEF, respondents are asked to rate the frequency of the child or adolescent's behavior such as their ability to finish tasks or complete multistep instructions. Behaviors were rated on a scale from 1 to 3, where 1 = never, 2 = sometimes, and 3 = often; more frequent failures of WM resulted in higher raw scores. In an initial test review, Gioia et al. (2000) reported satisfactory reliability for the BRIEF. Indeed, internal consistency as measured by Chronbach α ranged from .80 to .98 for clinical and normative samples. Test–retest reliability for the parent form was also strong (r = .81). The primary rating-based measure used in the present study was the raw score obtained by summing across the 10 items of the BRIEF–WM subscale. Total scores ranged from 10 to 30 for each participant.

In addition, primary caregivers also rated the frequency with which their child exhibited each of the 18 ADHD symptoms from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM–IV; American Psychiatric Association, 1994) using the DuPaul ADHD Rating Scale (DuPaul et al., 1998). Symptoms were rated on a scale from 0 to 3, where 0 = never or rarely, 1 = sometimes, 2 = often, and 3 = very often. Caregiver ratings were tallied within and across the inattentive and hyperactive/impulsive symptom domains. The DuPaul scale has been widely used to measure the number and severity of ADHD symptoms, and the psychometric properties of this measure appear to be adequate, as demonstrated by good internal reliability coefficients, high test–retest reliability, and effective discriminatory power (DuPaul et al., 1998). Total symptom scores ranged from 0 to 54 for each participant.

Participant performance-based measures

Participants were administered the WASI and the TOWRE to estimate full-scale IQ and word reading efficiency, respectively. These scales were included in the original study to measure cognitive abilities that may be confounded with ADHD diagnostic status; however, Gibson et al. (2018) showed that these abilities did not contribute to group differences in WM, and they are not discussed further in the present study. The performance-based estimates of PM and SM capacities were assessed by verbal and spatial IFR tasks (Gibson et al., 2010). IFR tasks were considered appropriate measures of WM based on research by Unsworth and Engle (2007a) demonstrating that a verbal IFR task loaded as highly as did complex span WM tasks on a latent WM construct. Further, traditional complex and simple span tasks do not permit analysis of the distinct contributions of PM and SM, sometimes even conflating those components, whereas detailed analysis of IFR task performance can be conducted by inspecting the number of items correctly recalled from both PM and SM, as well as serial position effects.

Following Gibson et al. (2010), the verbal IFR task included 15 lists of 12 unique high-frequency words presented in random order. All participants saw the same word lists presented on a standard cathode ray tube (CRT) monitor in 20-point white font on a black background. Each word was presented consecutively in the middle of the computer screen for 1 s followed by a 1-s interword delay. Following presentation of the entire list, question marks appeared in the center of the screen signifying the recall period.

Importantly, a participant's order of report strategy can influence estimates of PM and SM capacities (Gibson et al., 2011). Thus, participants had 30 s to recall aloud as many words as possible with the explicit instruction to start recalling items from the end of the presentation list first, although items need not be recalled in consecutive order. Providing such instructions has shown to induce a recency strategy, which is ideal for ensuring that the same recall strategy is used across participants and for estimating PM and SM capacities (Craik & Birtwistle, 1971). Further, participants needed to wait for the entire 30 s to elapse before beginning the next trial so that they could not terminate their recall early if they found the delay during the recall period aversive (Sonuga-Barke, 2003). Prior to the test trials, three practice trials using letter stimuli were presented. Participant responses were recorded via a microphone and digital audio recorder and scored as correct if they matched a word on the list, if they uttered a plural version of a list item that was singular, or if they said a past-tense version of a list item that was in the present tense.

For the spatial IFR task, participants were shown 15 lists of 12 different spatial locations marked by white squares on a black background of a standard CRT monitor. Each of the 12 locations per list were cued by temporarily changing color from white to red for 1 s followed by a 1-s delay before the next location was cued. Following presentation of the entire list, question marks appeared in the center of the screen signifying the recall period. Participants had 30 s to click as many of the cued locations as possible with the explicit instruction to begin recalling at the end of the list first, although items need not be clicked in consecutive order. Again participants were required to wait for the recall period to fully elapse before proceeding to the next trial, and only then did the test administrator initiate the next trial.

Squares appeared at any one of 180 (15 lists × 12 locations) unique locations on the computer screen, and each location was cued only once across the 15 different lists. To make the task more manageable, 36 of the possible 180 squares were randomly selected and appeared in three consecutive trials (3 trials × 12 cued locations = 36 locations). After the first three trials, 36 new locations were randomly selected from the possible 180 without replacement, and those appeared in the last three trials. Prior to the test trials, three practice trials were presented. Participant responses were recorded by the computer and included the location of the mouse clicks, the order of the mouse clicks, and the number of correct responses. Items were scored as correct if they matched one of the presented list items from the current trial.

Scoring participant performance-based measures

Tulving and Colotla's (1970) method was used to estimate the number of items correctly recalled from PM and SM during the verbal and spatial IFR tasks. In applying the method to IFR task performance, estimates of PM and SM capacities take into account both the amount of input and output interference. The greater the amount of interference preceding the recall of an item, the more likely an item was recalled from SM rather than PM. Thus, for an item to be recalled from PM, there must be seven or fewer items between presentation and recall. For an item to be recalled from SM, there must be more than seven items between presentation and recall.

For instance, suppose a participant was presented with a 12-item list of letters such as L, M, N, O, P, Q, R, S, T, U, V, W and then recalled W, V, T, L, N, P. In this case, three items were recalled from PM: W (0 items intervened between presentation and recall), V (two items intervened), and T (five items intervened). The remaining letters were recalled from SM: L (14 items intervened), N (13 items intervened), and P (12 items intervened). Such estimates of PM and SM capacities have been validated by other researchers in demonstrating that SM capacity but not PM capacity is affected by the build-up of proactive interference (Craik & Birtwistle, 1971; Unsworth & Engle, 2007a; Unsworth, Spillers, & Brewer, 2010). Using the Tulving and Colotla (1970) method, the number of items recalled from each PM and SM component were averaged across the 15 trials from each task modality, resulting in estimates of verbal PM (VPM), verbal SM (VSM), spatial PM (SPM), and spatial SM (SSM) capacities.

Order of report strategy: Probability of first and correct recall

Participants were included in the final analysis if they used a recency order of report strategy during the IFR tasks, as indicated by probability of first recall and/or probability of correct recall functions (Unsworth, Brewer, & Spillers, 2011). When participants use a recency order of report strategy, they tend to begin recalling items from the end of list first (higher probability of first recall on recency items) and tend to correctly recall more items from the end of the list (higher probability of correct recall on recency items). However, when participants use a primacy recall initiation strategy, they tend to begin recalling items from the start of the list and correctly recall more items from the start of the list (Unsworth et al., 2011).

Probability of first recall is the number of times an item at each serial position in the presentation list was recalled first divided by the total number of trials. For instance, say a presentation list had items: L, M, N, O, P, Q, R, S, T, U, V, W. If a participant recalled item W first on five out of six trials, then the probability of first recall for item W was 5/6. If item V was recalled first on the remaining trial, the probability of first recall for item V was 1/6, and the probability of first recall for all remaining items was 0/6.

Following Unsworth et al. (2011), participant probability of first recall was averaged for the last three serial positions (e.g., items W, V, U) and first three serial positions (e.g., items L, M, N), and a difference score was calculated from those two averages. A participant was classified as having used a recency order of report strategy if the difference between average probabilities of first recall was .10 or greater (see Footnote 1 in Unsworth et al., 2011). Thus, in our example, average probability of first recall at the last three serial positions was .17 [(5/6 + 1/6 + 0/6)/6] and 0 [(0/6 + 0/6 + 0/6)/6] for the first three serial positions. The difference between these averages was .17. So, in this example a participant would have evidenced using a recency order of report strategy on the probability of first recall measure.

Probability of correct recall is the number of times an item at each serial position in the presentation list was recalled correctly divided by the total number of trials. Following Unsworth et al. (2011), participant probability of correct recall was averaged for the last three serial positions (e.g., items W, V, U) and first three serial positions (e.g., items L, M, N), and a difference score was calculated from those two averages. A participant was classified as having used a recency order of report strategy if the difference between average probabilities of correct recall was .10 or greater (see Footnote 1 in Unsworth et al., 2011).

For instance, returning to our presentation list: L, M, N, O, P, Q, R, S, T, U, V, W, suppose a participant recalled W correctly on six trials (6/6), item V correctly on four trials (4/6), item U correctly on five trials (5/6), item L correctly on three trials (3/6), item M on three trials (3/6), and item N on two trials (2/6). Then, average probability of correct recall for the last three serial positions was .42 [(6/6 + 4/6 + 5/6)/6] and .22 for the first three serial positions [(3/6 + 3/6 + 2/6)/6]. Given that the difference between the averages was .20, this participant would have also evidenced a recency order of report strategy based on the probability of correct recall measure.

Although probability of first recall and probability of correct recall functions each provide evidence for whether or not a recency order of report strategy was in fact employed, these measures do not indicate how strongly that strategy was used (i.e., the strength of recency, SOR), particularly across verbal and spatial task modalities (Gibson & Gondoli, 2017). Thus, a composite verbal strength of recency (VSOR) score was created by averaging the standardized difference scores from the probability of first and correct recall measures for the verbal IFR task. Similarly, a composite spatial strength of recency (SSOR) score was created using the probability of first and correct recall measures from the spatial IFR task. Higher scores on the VSOR and SSOR indicated a stronger recency strategy, where a higher number of items were recalled from PM, and a lower number of items were recalled from SM. Lower VSOR and SSOR scores indicated a weak recency order of report strategy, where a lower number of items were recalled from PM, and a high number of items were recalled from SM (Gibson et al., 2018).

Taken together, it is crucial to account for participant order of report strategy used during IFR tasks to ensure that the same strategy is used across participants and so that estimates of PM and SM capacities are not inflated or deflated (Unsworth et al., 2011). Perhaps more importantly, Gibson and Gondoli (2017) have shown that SOR across both verbal and spatial domains (i.e., VSOR and SSOR) operate as a type of third variable that suppresses the relation between PM and SM; that is, the uncontrolled operation of SOR causes the relation between PM and SM to be more negative than it otherwise would be, thereby suppressing the positive relation that naturally exists between these two components of WM. Hence, steps have been taken to control for SOR in this and previous studies.

Results

Of the 100 participants who completed the IFR tasks, eight (n = 6 with ADHD) were excluded because their responses on the verbal IFR task could not be retrieved from the tape recorders, and one participant was excluded because her SSM score was three standard deviations above the overall mean. An additional 17 participants (n = 10 with ADHD) were excluded for failing to adhere to the IFR task instructions regarding use of recall strategies. Thus, 77 participants, ages 11–15 years (M = 12.55, SE = 0.11), were included in the final analysis (n = 32 with ADHD, n = 45 without ADHD).

Thirty-two participants had a confirmed diagnosis of ADHD (n = 24 males), and 45 participants did not meet criteria for an ADHD diagnosis (n = 15 males). As reported in our prior article (Gibson et al., 2018), ADHD and control participants differed significantly on both verbal and spatial SM capacity, and they also differed significantly on verbal PM capacity (but not spatial PM capacity). Note that these group means were adjusted by using corresponding measures of SOR as a covariate in analysis of covariance (ANCOVA; see Table 1 in Gibson et al., 2018). Group differences in parent-rated BRIEF–WM scores were not explicitly reported in this prior article; however, as expected, parents reported a significantly greater number of WM-related problems in the ADHD group (M = 23.91, SE = 0.90) than in the control group (M = 13.53, SE = 0.78), $t(75) = 8.68$, $p < .001$.

Prior to estimating the correlations between the PM and SM components of WM and BRIEF–WM, it was first crucial to account for individual differences in the extent to which participants used the recency order of report strategy during the verbal and spatial IFR tasks. Accordingly, estimates of VPM capacity and VSM capacity were each regressed on estimates of VSOR, and unstandardized residuals were calculated by subtracting predicted VPM and VSM scores from their corresponding actual scores; these residuals were termed VPM_r and VSM_r. Similarly, estimates of SPM capacity and SSM capacity were each regressed on estimates of SSOR, and unstandardized residuals were calculated; these residuals were termed SPM_r and SSM_r.

Descriptive statistics for VPM_r, VSM_r, SPM_r, SSM_r, and BRIEF–WM, based on the total sample, are listed in Table 1. Note that measures of skewness and kurtosis were in the acceptable range for each of the variables. The lower diagonal of Table 2 lists the bivariate correlations between VPM_r, VSM_r, SPM_r, SSM_r, and BRIEF–WM. There was a significant negative correlation between each of the VPM_r, VSM_r, and SSM_r capacities and BRIEF–WM scores; the negative correlation between SPM_r and BRIEF–WM did not attain significance. Partial correlations were computed given that there was a significant positive bivariate correlation observed between estimates of VPM_r and VSM_r capacities and between estimates of SPM_r and SSM_r capacities. The upper diagonal of Table 2 lists the partial correlations between BRIEF–WM and each of the other performance-based measures of WM. Analysis of the verbal domain revealed that the correlation between VSM_r and BRIEF–WM scores ($r = -.45$, $p < .0001$) was significant when VPM_r was controlled;

Table 1. Descriptive statistics for the main variables in this study.

	M	SD	Skewness	Kurtosis
BRIEF–WM	17.84	7.27	0.33	−1.43
VPM_r	0.00	0.42	−0.16	−0.54
VSM_r	0.00	0.82	0.60	0.43
SPM_r	0.00	0.40	0.18	−0.72
SSM_r	0.00	0.86	0.06	−0.45

Note. Total: N = 77. BRIEF–WM = parent-rated Behavior Rating Inventory of Executive Function–Working Memory Scale; VPM_r = unstandardized residual of primary memory (PM) capacity estimated from the verbal immediate free recall (IFR) task; VSM_r = unstandardized residual of secondary memory (SM) capacity estimated from the verbal IFR task; SPM_r = unstandardized residual of PM capacity estimated from the spatial IFR task; SSM_r = unstandardized residual of SM capacity estimated from the spatial IFR task.

Table 2. Bivariate and partial correlations for the main variables in this study.

	1	2	3	4	5
1. BRIEF–WM	—	−.12	−.45**	−.08	−.24*
2. VPM_r	−.34**	—			
3. VSM_r	−.53**	.47**	—		
4. SPM_r	−.19	.46**	.22	—	
5. SSM_r	−.29*	.26*	.42**	.40**	—

Note. Bivariate correlations are reported below the diagonal, and partial correlations are reported above the diagonal. BRIEF–WM = parent-rated Behavior Rating Inventory of Executive Function–Working Memory Scale; VPM_r = unstandardized residual of primary memory (PM) capacity estimated from the verbal immediate free recall (IFR) task; VSM_r = unstandardized residual of secondary memory (SM) capacity estimated from the verbal IFR task; SPM_r = unstandardized residual of PM capacity estimated from the spatial IFR task; SSM_r = unstandardized residual of SM capacity estimated from the spatial IFR task.
*$p < .05$. **$p < .01$.

however, the correlation between VPM_r and BRIEF–WM scores ($r = -.12$, $p = .31$) was not significant when VSM_r was controlled. Likewise, analysis of scores from the spatial domain revealed a similar pattern to that for the verbal domain: The correlation between SSM_r and BRIEF–WM scores ($r = -.24$, $p = .04$) was significant when SPM_r was controlled; however, the correlation between SPM_r and BRIEF–WM scores ($r = -.08$, $p = .48$) was not significant when SSM_r was controlled. These findings support the main hypothesis that the partial correlation between BRIEF–WM scores and SM scores (controlling for PM) is more robust than the correlation between BRIEF–WM scores and PM scores (controlling for SM).

Examination of the observed bivariate and partial correlations listed in Table 2 also revealed an unexpected finding: namely, that the correlations between the two verbal estimates of WM and BRIEF–WM were numerically larger than the correlations between the two spatial estimates of WM and BRIEF–WM. Because these comparisons involved within-component comparisons, tests of the difference between dependent bivariate correlations were conducted to examine whether the verbal versus spatial correlations with BRIEF–WM differed significantly within each component of WM (Steiger, 1980). These tests indicated that the negative bivariate correlation between VSM_r and BRIEF–WM ($r = -.53$) was significantly larger than the negative bivariate correlation between SSM_r and BRIEF–WM ($r = -.29$), $z = 2.21$, $p = .03$; however, although numerically larger, the negative bivariate correlation between VPM_r and BRIEF–WM ($r = -.39$) did not differ significantly from the negative bivariate correlation between SPM_r and BRIEF–WM ($r = -.19$), $z = 1.31$, $p = .19$. The same pattern was also observed with partial correlations, but tests of the difference between dependent correlations could not be performed on the partial correlations. The significant difference observed between VSM_r and SSM_r is discussed in greater detail below.

Finally, we also examined the possibility that the significant relation observed between estimates of SM capacity and BRIEF–WM scores was moderated by total ADHD symptoms within each task modality. Multiple regression was conducted with two first-order predictors: total ADHD symptoms (DuPaul et al., 1998) and Residual SM capacity. Both first-order predictors were centered on their respective means (Aiken & West, 1991). More importantly, one second-order predictor was included to assess moderation: Total ADHD Symptoms × Residual SM Capacity. BRIEF–WM scores served as the dependent variable. With respect to the verbal modality, the overall model was statistically significant, $F(3, 73) = 35.03$, $p < .0001$, with the three predictors explaining approximately 59% of the variance ($R^2 = .59$, adjusted $R^2 = .57$). However, the critical interaction between total ADHD symptoms and VSM_r capacity scores did not approach significance ($p = .30$), suggesting that the total number of ADHD symptoms did not moderate the relationship between VSM_r and BRIEF–WM (see Table 3). Likewise, with respect to the spatial modality, the overall model was statistically significant, $F(3, 73) = 31.57$, $p < .0001$, with the three predictors explaining approximately 57% of the variance ($R^2 = .57$, adjusted $R^2 = .55$). However, the critical interaction between total ADHD symptoms and SSM_r capacity scores did not approach significance ($p = .64$), once again suggesting that the total number of ADHD symptoms did not moderate the relationship between SSM_r and BRIEF–WM (see Table 4).

Discussion

The current study was motivated by previous studies suggesting that performance- and rating-based measures of WM correlate poorly (Anderson et al., 2002; Mahone et al., 2009; Mahone et al., 2002; McAuley et al., 2010; Miranda et al., 2015; Toplak et al., 2009, 2013; Vriezen & Pigott, 2002). Some researchers have

Table 3. Regression coefficients and 95% confidence intervals to examine potential moderation by ADHD in the verbal modality.

Predictor	*B*	SE_B	β	*p*	95% CI Lower limit	95% CI Upper limit
ADHD	0.30	0.48	0.60	.00	0.20	0.40
VSM_r	−2.01	0.82	−0.23	.02	−3.64	−0.38
ADHD × VSM_r	−0.06	0.06	−0.09	.30	−0.17	0.05
Total R^2	.59					
Adjusted R^2	.57					

Note. ADHD = parent-rated total attention-deficit/hyperactivity disorder symptoms; VSM_r = unstandardized residual of secondary memory (SM) capacity estimated from the verbal immediate free recall (IFR) task; CI = confidence interval. Both ADHD and VSM_r were centered on their respective means.

Table 4. Regression coefficients and 95% confidence intervals to examine potential moderation by ADHD in the spatial modality.

Predictor	*B*	SE_B	β	*p*	95% CI Lower limit	95% CI Upper limit
ADHD	0.36	0.04	0.71	.00	0.28	0.44
SSM_r	−0.84	0.69	−0.10	.23	−2.21	0.54
ADHD × SSM_r	−0.02	0.05	−0.04	.64	−0.11	0.07
Total R^2	.57					
Adjusted R^2	.55					

Note. ADHD = parent-rated total attention-deficit/hyperactivity disorder symptoms; SSM_r = unstandardized residual of secondary memory (SM) capacity estimated from the spatial immediate free recall (IFR) task; CI = confidence interval. Both ADHD and SSM_r were centered on their respective means.

hypothesized that the reason these measures fail to converge is that performance-based measures assess how efficiently participants engage information-processing mechanisms under highly controlled conditions where the task goal is known, whereas rating-based measures assess how well executive processes can be engaged when the task goal must be initiated and executed independently by the participant (Barkley & Fischer, 2011; Ten Eycke & Dewey, 2016; Toplak et al., 2013). Based on this interpretation, performance- and rating-based WM measures assess different constructs, and, thus, both measure types would be needed to capture distinct information about participant WM deficits or lack thereof.

In the present study, we hypothesized that performance- and rating-based WM measures may converge, to a larger extent, when performance-based measures are theoretically motivated and account for the unique contributions of both PM and SM components of WM. That is to say, the previously observed poor relation between measure types may be due to use of tasks that assess solely PM at shorter list lengths or a mixture of PM and SM at supralist lengths (i.e., digit span or spatial span; Toplak et al., 2013; Unsworth & Engle, 2006). In essence then, these studies may have been procedurally controlling for the distinct variance associated with SM capacity much as the current study statistically controlled for the influence of SM capacity on PM capacity when estimating the partial correlation between PM and BRIEF–WM.

More specifically, there was a significant negative correlation observed between VSM_r capacity and BRIEF–WM scores when VPM_r capacity was controlled. And, likewise, there was a significant negative correlation observed between SSM_r capacity and BRIEF–WM scores when SPM_r capacity was controlled. However, although there was a significant bivariate correlation between VPM_r capacity and BRIEF–WM scores in the current study, we showed that this negative correlation was driven by the overlap between VPM_r and VSM_r capacities, as VPM_r capacity did not significantly relate to BRIEF–WM scores when VSM_r capacity was controlled. And, likewise, SPM_r capacity did not significantly relate to BRIEF–WM scores when SSM_r capacity was controlled.

In our view, the negative correlation observed between performance-based measures of WM capacity and ratings-based measures of WM-related problems was found to be more robust for estimates of SM capacity than for estimates of PM capacity because weaknesses in SM capacity are more likely to lead to problems in daily living than weaknesses in PM capacity. For instance, consider the following anecdotal situation: You are sitting at the kitchen table reading the morning news when you decide to get a glass of orange juice from the refrigerator. You get up from the table, make your way across the room, and open the door to the refrigerator. But, as you stand there with the door open, you realize that you are not entirely sure why. At that very moment, as you stand there with no goal-relevant information in the forefront of your mind, you may be distracted by or act impulsively toward other salient external objects or internal thoughts. You may end up fixing a scrambled egg or drinking a glass of chocolate milk instead.

In this scenario, the failure to maintain the “get orange juice” goal clearly reflects a limitation in the PM component of WM capacity. However, although the failure to maintain goal-relevant information in PM may be a necessary condition for the temporary disruption of goal-directed behavior, it is not sufficient for the complete termination of that behavior. This is because goal-relevant information that has been displaced from PM may still be recovered from SM using cue-dependent retrieval mechanisms (Mogle et al., 2008; Unsworth, 2007, 2009; Unsworth & Engle, 2007a, 2007b; Unsworth & Spillers, 2010). In other words, although you may have temporarily lost the critical goal-relevant information from PM by the time you showed up at the refrigerator door, you may have been able to resume the intended goal-directed behavior by retrieval of the proper goal-relevant information from SM back into PM. Thus, if this analysis is correct, the extent to which goal-directed behaviors are terminated prematurely and lead to a WM-related problem should be more strongly associated with weaknesses in SM capacity (holding PM capacity constant) than with weaknesses in PM capacity (holding SM capacity constant).

However, the notion that weaknesses in SM capacity may be more predictive of WM-related problems than PM capacity must also be qualified by the present finding that the relation between SM and BRIEF–WM was found to be significantly larger in the verbal modality than in the spatial modality. One intriguing explanation for this difference concerns the possibility that free recall during the verbal IFR task was less structured than free recall during the spatial IFR task.

More specifically, with respect to retrieval from SM, participants had to use cues to define a memory search set from which potential list items were retrieved (Unsworth, 2007, 2009). In the verbal task, the generation of these cues would be mainly controlled by the participant because each verbal item in the list was terminated after one second, and these items were not visually present during the recall portion of the task.

However, in the spatial task, the generation of these cues was at least partially controlled by the environment because, although each spatial item in the list was also terminated after one second in the spatial task, the set of potential locations (i.e., relevant and irrelevant boxes on the computer screen) remained visually present during the recall portion of the task. In other words, in the context of the anecdotal refrigerator scenario described above, cue-dependent retrieval from the verbal IFR task was more analogous to retrieving the lost "get orange juice" goal with the refrigerator door closed (a harder task), whereas retrieval from the spatial IFR task was more analogous to retrieving the lost goal with the refrigerator door open (an easier task). In this way, the correlation between VSM_r and BRIEF–WM may have been larger than the correlation between SSM_r and BRIEF–WM because VSM_r is a better measure of how well executive processes can be engaged when the task goal must be initiated and executed independently by the participant (Toplak et al., 2013). However, stronger evidence for this hypothesis awaits a manipulation of internally versus externally retrieval cues that is not confounded with task modality.

It is worth noting that the current results, although preliminary, may have implications for assessment and intervention. Regarding assessment, researchers and clinicians alike may benefit from an increased understanding of WM as a complex ability that requires measurement of both PM and SM capacities, as well as rating-based measures to fully explain the nature of WM deficits or lack thereof. Furthermore, the relationship between performance- and rating-based measures might be further elucidated by the use of composite scores or factor analytic techniques that account for the common variance among numerous WM tasks while limiting the task-specific variance associated with using a limited number of WM tasks (Schreiber, Possin, Girard, & Rey-Casserly, 2014; Snyder, Miyake, & Hankin, 2015). For instance, including delayed free recall tasks to measure SM retrieval mechanisms (i.e., search set size, recovery, and error monitoring; Unsworth, 2009) and/or complex span tasks to isolate SM, as well as tasks that capture specific aspects of PM ability (i.e., capacity and controlled attention; Unsworth, Fukuda, Awh, & Vogel, 2014) may allow for a more nuanced examination of how rating-based WM measures relate to specific processes or mechanisms underlying both PM and SM capacities.

Regarding interventions, the current study suggested that SM capacity might be an important target for remediation of WM deficits. Interestingly, there has been a recent proliferation of computerized cognitive training paradigms aimed at improving WM (Melby-Lervåg, Redick, & Hulme, 2016). At the core of this method is the idea that sustained and intense practice on span tasks will strengthen the neural substrates underlying WM and lead to enhanced performance on WM-related abilities in daily life (e.g., far transfer to on-task behavior; Klingberg, 2010). However, research from our group has indicated that extant training programs primarily target PM capacity (Gibson et al., 2011). Furthermore, recent attempts to target SM have sometimes (Gibson et al., 2013; Harrison et al., 2013), but not always (Gibson et al., 2011, 2012; Ralph et al., 2017), enhanced SM capacity scores. Thus, because extant paradigms have yet to reliably target both WM components, development of novel training regimens targeting SM capacity still remains an open area for investigation.

To summarize, the unique relationship between SM capacity and parent-rated BRIEF–WM scores implies that earlier studies have not adequately considered the complex nature of the WM construct and the need for traditional performance-based WM measures to account for both PM and SM components. Future work should evaluate whether rating-based WM measures and, by extension, real-world WM-related behaviors can be predicted with performance-based measures that account for distinct components of WM, as well as determine that both measure types converge at the neuroanatomical level. Clarifying the construct validity of measurement tools by aligning performance- and rating-based measures with modern cognitive theory will be a crucial step if researchers and clinicians aim to appropriately identify deficit executive functions and target those abilities for intervention.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Kathryn J. Ralph http://orcid.org/0000-0002-2127-0742

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. London, England: Sage.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM)* (4th ed.). Washington, DC: Author.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology*, *8*, 231–240.

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology*, *36*, 137–161.

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

Craik, F. I. M., & Birtwistle, J. (1971). Proactive inhibition in free recall. *Journal of Experimental Psychology*, *91*, 120–123.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD rating scale—IV: Checklist, norms, and clinical interpretation*. New York, NY: The 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*, *128*, 309–331.

Faridi, N., Karama, S., Burgaleta, M., White, M. T., Evans, A. C., Fonov, V., … Waber, D. P. (2015). Neuroanatomical correlates of behavioral rating versus performance measures of working memory in typically developing children and adolescents. *Neuropsychology*, *29*, 82–91.

Gibson, B. S., & Gondoli, D. M. (2017). Unmasking the component-general and component-specific aspects of primary and secondary memory in the immediate free recall task. *Memory & Cognition*. Advance online publication. doi:10.3758/s13421-017-0769-1

Gibson, B. S., Gondoli, D. M., Flies, A. C., Dobrzenski, B. A., & Unsworth, N. (2010). Application of the dual-component model of working memory to ADHD. *Child Neuropsychology*, *16*, 60–79.

Gibson, B. S., Gondoli, D. M., Johnson, A. C., Steeger, C. M., Dobrzenski, B. A., & Morrissey, R. A. (2011). Component analysis of verbal versus spatial working memory training in adolescents with ADHD: A randomized, controlled trial. *Child Neuropsychology*, *17*, 546–563.

Gibson, B. S., Gondoli, D. M., Kronenberger, W. G., Johnson, A. C., Steeger, C. M., & Morrissey, R. A. (2013). Exploration of an adaptive training regimen that can target the secondary memory component of working memory capacity. *Memory & Cognition*, *41*, 726–737.

Gibson, B. S., Gondoli, D. M., Ralph, K. J., & Sztybel, P. (2018). Application of the dual-component model of working memory to ADHD: Greater secondary memory deficit despite confounded cognitive differences. *Child Neuropsychology*, *24*, 61–81.

Gibson, B. S., Kronenberger, W. G., Gondoli, D. M., Johnson, A. C., Morrissey, R. A., & Steeger, C. M. (2012). Component analysis of simple span vs. complex span adaptive working memory exercises: A randomized, controlled trial. *Journal of Applied Research in Memory and Cognition*, *1*, 179–184.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function*. Odessa, FL: Psychological Assessment Resources.

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Epsy, K. A. (2002). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF) in a clinical sample. *Child Neuropsychology*, *8*, 249–257.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, *24*, 2409–2419.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*, 637–671.

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

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*, 317–324.

Mahone, E. M., Martin, R., Kates, W. R., Hay, T., & Horská, A. (2009). Neuroimaging correlates of parent ratings of working memory in typically developing children. *Journal of the International Neuropsychological Society*, *15*, 31–41.

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., & Denckla, M. N. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or tourette syndrome. *Archives of Clinical Neuropsychology*, *7*, 643–662.

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*, 495–505.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer" evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*, 512–534.

Miranda, A., Colomer, C., Mercader, M., Fernández, M. I., & Presentación, M. J. (2015). Performance-based versus behavioral ratings in the assessment of executive functioning in preschoolers: Associations with ADHD symptoms and reading achievement. *Frontiers in Psychology*, *6*, 545.

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological Science*, *19*, 1071–1077.

Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsycholgists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology*, *31*, 206–230.

Ralph, K. J., Gibson, B. S., Gondoli, D. M., Sztybel, P., Pauszek, J. R., Miller, R. W., & Litzow, E. (2017). Targeting the three stages of retrieval from secondary memory in a double-blinded, placebo-controlled, randomized working memory training study. *Journal of Cognitive Enhancement*, *1*, 455–477.

Schreiber, J. E., Possin, K. L., Girard, J. M., & Rey-Casserly, C. (2014). Executive function in children with attention deficit/hyperactivity disorder: The NIH EXAMINER battery. *Journal of the International Neuropsychological Society*, *20*, 41–51.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH diagnostic interview schedule for children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 28–38.

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*, 328.

Sonuga-Barke, E. J. S. (2003). The dual pathway model of AD/HD: An elaboration of neuro- developmental characteristics. *Neuroscience & Biobehavioral Reviews*, *27*, 593–604.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251.

Ten Eycke, K. D., & Dewey, D. (2016). Parent-report and performance-based measures of executive function assess different constructs. *Child Neuropsychology*, *22*, 889–906.

The Psychological Corporation. (1999). *Manual for the Wechsler abbreviated scale of intelligence*. San Antonio, TX: Author.

Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2009). Executive functions: Performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*, 53–72.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*, 131–143.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of reading word efficiency*. Austin, TX: Pro-ed.

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology*, *1*, 86–98.

Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 1020–1034.

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition*, *37*, 837–849.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2011). Inter- and intra-individual variation in immediate free recall: An examination of serial position functions and recall initiation strategies. *Memory*, *19*, 67–82.

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*, 68–80.

Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132.

Unsworth, N., & Engle, R. W. (2007b). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*, 1038–1066.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26.

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attentional control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*, 392–406.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2010). The contributions of primary and secondary memory to working memory capacity: An individual differences analysis of immediate free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 240–247.

Vriezen, E. R., & Pigott, S. E. (2002). The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology*, *8*, 296–303.