

# **Journal of Experimental Psychology: Learning, Memory, and Cognition**

## **The Role of Maintenance and Disengagement in Predicting Reading Comprehension and Vocabulary Learning**

Jessie D. Martin, Zach Shipstead, Tyler L. Harrison, Thomas S. Redick, Michael Bunting, and Randall W. Engle

Online First Publication, June 6, 2019. <http://dx.doi.org/10.1037/xlm0000705>

### **CITATION**

Martin, J. D., Shipstead, Z., Harrison, T. L., Redick, T. S., Bunting, M., & Engle, R. W. (2019, June 6). The Role of Maintenance and Disengagement in Predicting Reading Comprehension and Vocabulary Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <http://dx.doi.org/10.1037/xlm0000705>

# The Role of Maintenance and Disengagement in Predicting Reading Comprehension and Vocabulary Learning

Jessie D. Martin  
Georgia Institute of Technology

Zach Shipstead  
Colby College

Tyler L. Harrison  
University of North Georgia, Dahlonega, Georgia

Thomas S. Redick  
Purdue University

Michael Bunting  
University of Maryland, College Park, Maryland

Randall W. Engle  
Georgia Institute of Technology

This study uses a novel framework based on work by Shipstead, Harrison, and Engle (2016) that includes measures of both working memory capacity and fluid intelligence in an attempt to better understand the processes that influence successful reading comprehension at the latent level. Further, we extend this framework to a second educationally relevant ability: second-language vocabulary learning. A large sample of young adults received a battery of working memory, fluid intelligence, language comprehension, and memory updating tasks. The results indicate that individual differences in reading comprehension and vocabulary learning benefit from the ability to maintain active information, as well as to disengage from no longer relevant information. Subsequently, we provide an interpretation of our results based on the maintenance and disengagement framework proposed by Shipstead et al. (2016).

**Keywords:** maintenance, disengagement, working memory capacity, fluid intelligence, reading comprehension

Despite decades of research examining processes necessary for successful reading comprehension, we still do not fully understand or agree on the cognitive construct which underlie performance. For example, we have established that working memory capacity is important for predicting reading comprehension performance (Daneman & Carpenter, 1980); however, a substantial amount of variance remains unaccounted for in studies exploring the relationship between working memory capacity and reading comprehension (Borella, Carretti, & Pelegrina, 2010; Chiappe, Siegel, & Hasher, 2000; Christopher et al., 2012; Engle, Cantor, & Carullo, 1992; Was & Woltz, 2007). In an attempt to further explain other factors that are essential to performance, many of these

studies have begun to take an increasingly fractionated approach to filling in the pieces. This deconstructive approach has resulted in seemingly conflicting conclusions regarding mechanisms of interest, and has not left room for individual differences in mechanisms used in retrieval as well as organized forgetting. In an effort to return to a more parsimonious approach, we proposed a more process-general latent variable framework for the understanding of individual differences in reading comprehension.

Specifically, we contend that the current foundation for understanding reading comprehension has overemphasized the use of tasks which primarily reflect maintenance of information, while ignoring the beneficial role of forgetting—or what we will refer to as disengagement. We argue that individual differences in reading comprehension are not strictly related to maintenance processes captured by complex span measures of working memory capacity, but also to disengagement processes captured by measures of fluid intelligence. This approach is based on work by Shipstead et al. (2016) that defines the primary mechanisms of executive attention in terms of general processes of maintenance and disengagement (reflected in complex span and fluid intelligence measures, respectively), rather than through increasingly specified mechanistic functions like shifting, updating, and inhibition (Miyake et al., 2000; see Friedman & Miyake, 2017 for an updated perspective). We will identify advantages to our process general approach as well as limitations to current, more deconstructive, approaches including those that emphasize the role of processing speed (Christopher et al., 2012)

---

Jessie D. Martin, Department of Psychology, Georgia Institute of Technology; Zach Shipstead, Department of Psychology, Colby College; Tyler L. Harrison, Department of Psychology, University of North Georgia, Dahlonega, Georgia; Thomas S. Redick, Department of Psychology, Purdue University; Michael Bunting, Center for the Advanced Study of Language, University of Maryland, College Park, Maryland; Randall W. Engle, Department of Psychology, Georgia Institute of Technology.

This work was supported by grants from the Office of Naval Research (N00014-12-1-0406 and N00014-12-1-1011) to Randall W. Engle.

Correspondence concerning this article should be addressed to Jessie D. Martin, Department of Psychology, Georgia Institute of Technology, North Avenue NW, Atlanta, GA 30332. E-mail: [jessie.martin@gatech.edu](mailto:jessie.martin@gatech.edu)

and inhibition in reading comprehension (Hasher & Zacks, 1988; Zacks & Hasher, 1994).

### Reading Comprehension and Working Memory

Early theories of reading comprehension examined whether short-term memory (STM) capacity, or the amount of information an individual can maintain and recall correctly, predicted reading comprehension performance (Perfetti & Lesgold, 1977; Rizzo, 1939). The fundamental idea was that the more content individuals could hold in mind and later recall, the better. STM, as measured by “simple span” measures (serial recall tasks) were initially used to test this theory. While these serial recall measures did predict a degree of performance, the relationships were much smaller than anticipated (Daneman & Merikle, 1996). Rather, evidence pointed to a stronger relationship between reading comprehension and complex span measures, which are a more dynamic measure of working memory capacity (Daneman & Carpenter, 1980; Turner & Engle, 1989). For the purposes of this project, we define working memory capacity in terms of complex span measures. Whereas STM reflects a static memory capacity, working memory refers to a limited-capacity system where memory and attention interact to simultaneously store and process information in the service of complex cognition (Baddeley & Hitch, 1974). Complex span tasks combine storage and processing components, and reflect the attention driven processes of the central executive (Engle, 2002).

Daneman and Carpenter (1980) were the first to demonstrate that complex span tasks strongly predicted reading comprehension using the reading span task. Daneman and Merikle (1996), further showed that, complex span tasks were better predictors of reading comprehension than were simple span tasks because complex span tasks engage general processing plus storage measures. Additionally, the predictive value of complex span tasks was independent of their reliance on the manipulation of linguistic items (i.e., the reading span used by Daneman and Carpenter). These general findings have been repeatedly verified (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Budd, Whitney, & Turley, 1995; Engle et al., 1992; Turner & Engle, 1989).

More recently, focus has shifted toward isolating aspects of cognition which could further facilitate processing advantages in addition to working memory capacity. These processes include inhibition (Chiappe et al., 2000; Hasher & Zacks, 1988), processing speed (Christopher et al., 2012), and updating/general intelligence (Was, Rawson, Bailey, & Dunlosky, 2011). However, one construct which is highly related to working memory capacity that is notably absent from the study of reading comprehension is fluid intelligence.

### Maintenance and Disengagement as Working Memory Capacity and Fluid Intelligence

Fluid intelligence refers to the ability to solve novel problems, or to reason with novel information (Cattell, 1943; Horn & Cattell, 1966). Fluid intelligence and working memory capacity share a substantial amount of variance, with correlations ranging from .6–.9 (Kane, Hambrick, & Conway, 2005; Shipstead, Redick, Hicks, & Engle, 2012). According to the theory proposed by Shipstead et al. (2016) one potential explanation of the strong, but less than perfect relationship between working memory capacity and fluid intelligence is that the two constructs primarily reflect different processes, but both rely on domain-free executive attention for their implementation. Specifically, (complex span) measures of working memory capacity largely reflect the ability to maintain information while fluid intelligence measures reflect the need to disengage from no longer relevant stimuli (see Figure 1). This is not to say that the processes of maintenance and disengagement function in complete isolation across these task sets, but that each relies on one more than the other.

For example, on the Ravens progressive matrices, which is similar to the task shown in Figure 2, as individuals test new hypotheses within each trial they must continually test and replace the rejected hypothesized rule used to solve the pattern until the correct rule is discovered (Harrison, Shipstead, & Engle, 2015; Verguts & De Boeck, 2002; Wiley, Jarosz, Cushen, & Colflesh, 2011). This release of no-longer-relevant solution rules should also lead to that rejected rule being less likely to be retrieved immedi-

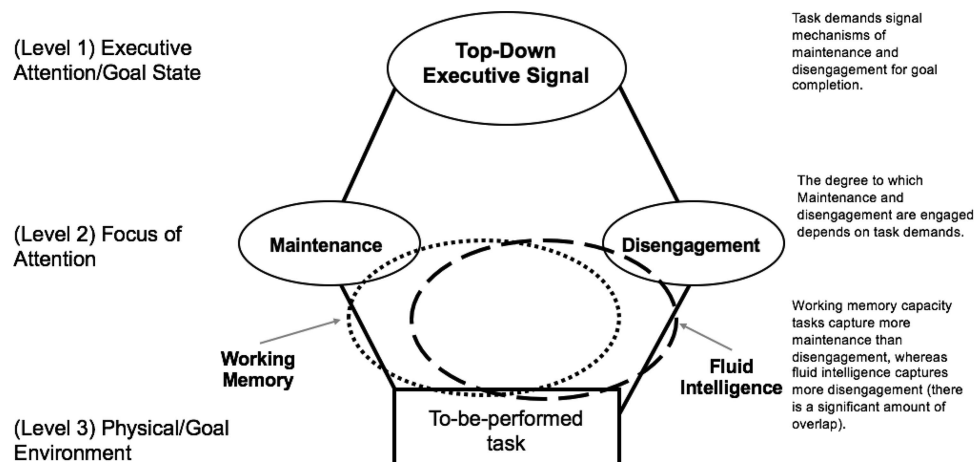


Figure 1. Maintenance and disengagement theory of working memory capacity adapted from Shipstead et al. (2016) with permission. The top-down signal organizes controlled maintenance and disengagement in the service of completing a specific goal.

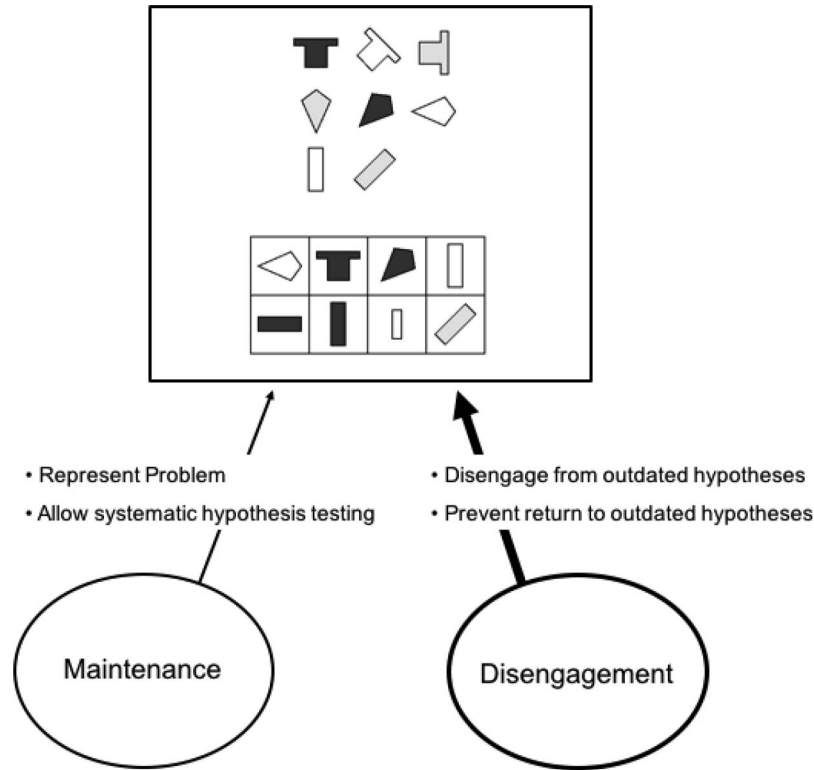


Figure 2. The proposed roles of maintenance and disengagement of a fluid intelligence task. Bolded circle indicates process of primary relevance. Adapted from Shipstead et al. (2016) with permission.

ately in a search for a new rule (Raaijmakers & Shiffrin, 1980). Further, failure to release or block this outdated information or hypothesis would result in participants re-retrieving the erroneous hypothesis and fixating, or perseverating. Fixating, perseverating, or retesting erroneous hypotheses all result in a lower score as the Ravens progressive matrices is scored based on the number of items completed in a given time. Although maintenance is present to the extent that individuals must maintain a hypothesis as it is being tested, we will later demonstrate how disengagement can be isolated through the use of structural equation modeling.

In contrast to fluid intelligence measures, working memory capacity measures are primarily maintenance oriented within each trial as target items must be maintained over a processing interval that includes interference. Between trials, it is beneficial for individuals to disengage from material presented in earlier trials, but this influence is minimal in comparison with the importance of maintain information (see Figure 3; see Kane, Bleckley, Conway, & Engle, 2001; Unsworth & Engle, 2006). This distinction between working memory capacity measures as relying primarily on maintenance and fluid intelligence tasks relying primarily on disengagement was illustrated by Shipstead et al. (2016) with a comparison of performance on an N-back task.

During this N-back task, individuals responded based on whether or not they observed the same stimulus (e.g., face) three trials ago. Lures were also included in this task design. Lures were defined as target face presented at a position other than three positions back (i.e., presented at two or four positions back). Lures close to the target are called “near” lures, while lures 8, 9, and 10 trials ago are

considered to be “far” lures. In theory, near lures should result in more false alarms than far lures because of their proximity to the target. However, as the distance between targets and lures increased, so did the correlation between this distance and measures of working memory capacity as well as fluid intelligence. Using partial correlation techniques, when working memory was controlled, the correlation between far lures and fluid intelligence was reduced but still increased as the distance between targets and lures increased. However, when fluid intelligence was controlled for, the relationship between working memory capacity and lure position was no longer significant. Thus, fluid intelligence but not working memory capacity predicted whether or not an individual would false alarm to a far lure. Shipstead et al. (2016) argued that the relationship between far lures and fluid intelligence reflected a failure to disengage from no-longer relevant information.

A second finding by Shipstead et al. (2016) further highlighted relationships that are driven by fluid intelligence that may otherwise appear to be related to working memory capacity. Rosen and Engle (1997) found that individuals who scored high on measures of working memory capacity outperformed low working memory capacity individuals on measures of verbal fluency. Additionally, low working memory capacity individuals, tended to repeat answers more than high capacity individuals. However, Shipstead et al. conducted a follow up to this study in which measures of fluid intelligence were also included. This time, the variance in verbal fluency was accounted for by fluid intelligence, but not working memory capacity.

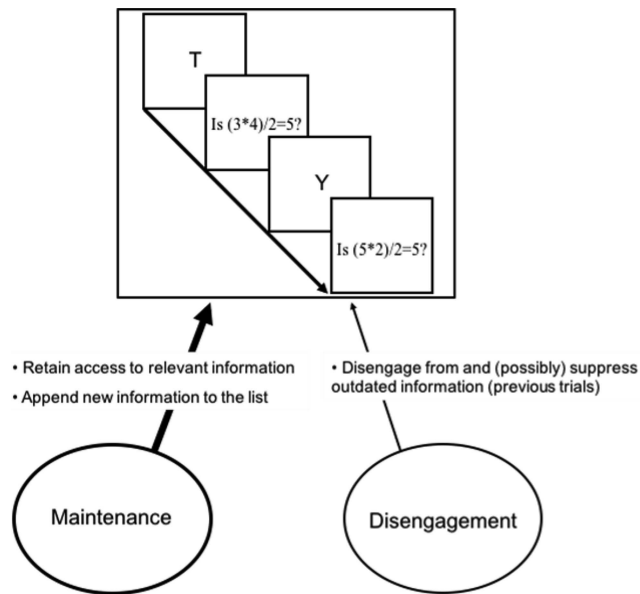


Figure 3. The proposed roles of maintenance and disengagement of a working memory capacity task. Bolded circle indicates process of primary relevance. Adapted from Shipstead, Harrison, and Engle (2016) with permission.

Rosen and Engle (1997) did not include measures of fluid intelligence, and also used an extreme-groups design. Subsequently, because of the shared attention-based variance between working memory capacity and fluid intelligence, the relationship between fluid intelligence and verbal fluency was overlooked. In essence, this is the same argument we make in this article regarding reading comprehension. Specifically, contributions related to fluid intelligence beyond working memory capacity likely exist, but have been obscured by the strong relationship between working memory capacity and fluid intelligence masking the role of fluid intelligence.

We do not go so far as to define all of the possible mechanisms of disengagement here, as our goal is to examine the process independent of the mechanism used to achieve it. Overall, disengagement reduces inappropriate retrieval and/or activation of no longer relevant information. This process may include the unbinding of items (Oberauer, 2005), tagging items as no longer relevant (Raaijmakers & Shiffrin, 1980), and substitution (Ecker, Lewandowsky, Oberauer, & Chee, 2010), and may be related to what Ecker, Lewandowsky, and Oberauer (2014) refer to as a distractor-removal process (see also Oberauer & Lewandowsky, 2016; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012). The distractor-removal process allows for the reduction of interference from irrelevant information, in turn reducing processing demands. Although this kind of disengagement may function at the between trial level of complex-span measures, we argue that these disengagement processes are more pronounced at the latent level with regard to fluid measures including fluid intelligence and updating. We do also wish to reemphasize that disengagement may be any or all of these types of mechanisms between and within individuals.

While reading, individuals need to not only maintain and add to information encountered previously, but they must also ensure that

representations are interpreted as intended in the context of the passage. For example, when reading words that have multiple meanings, such as “bat,” it is important to decide from the context whether bat refers to a baseball bat or a winged mammal. To ensure the proper interpretation is maintained, one must also release the activated alternative interpretation of the word bat, to reduce the likelihood that it is retrieved later. Additionally, action presented in one passage may later be negated in future contexts in such a way that continuing to maintain information from a previous passage or setting is not advantageous in understanding future context. Maintaining distinct timelines of events is one example in which both maintaining information active is advantageous, as well as being able to release no-longer-relevant information that might confuse the time course of events that have occurred. In both of these situations, complete and indiscriminate maintenance of information would result in excessive, and confusing representations for understanding as well as later recall.

### Why Introduce the Term Disengagement?

Disengagement is a general process of forgetting or releasing once, but no-longer-relevant information, and can encompass many mechanisms as listed previously. We use the term disengagement over specific mechanisms for several reasons. First, measuring domain-general forgetting through the use of fluid intelligence at the latent level allows for the use of a factor with strong and consistent loadings. Studies by Engle and colleagues frequently show factor loadings of .7 and above for all indicators of fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999; Shipstead et al., 2016) whereas factors of independent “executive” processes that have lower and less-consistent loadings (Miyake et al., 2000). As such, if there are domain-general contributions of forgetting, we are more likely to observe these relationships through the use of a more robust factor. Second, it is entirely possible that in complex cognitive tasks individuals are not uniform in their use of organized forgetting mechanisms. As such, an agnostic process-general term is preferable in describing the processes that underlie successful performance. Third, we view the disengagement perspective to be in line with findings regarding inhibition and reading comprehension (Borella et al., 2010; Christopher et al., 2012), but preferable because of its process-general nature.

For example, inhibition has been defined in many ways by many research groups, making it a difficult process to compare across studies. Some researchers define inhibition in terms of performance on other reading-based tasks (Borella et al., 2010), while others define inhibition in terms of response suppression measures (Christopher et al., 2012), while others still see these measures misleading or at best as falling under the umbrella of the common executive processes more broadly defined (Friedman & Miyake, 2017; Rey-Mermet, Gade, & Oberauer, 2018).

We wish to specifically highlight the definition of inhibition used by Christopher et al. (2012), as it is relevant to the interpretation of our results. Christopher et al. define inhibition in terms of an individual’s ability to ignore or suppress a prepotent response. These responses include avoiding attention capture by a flash in an antisaccade task measure, or responding to an incongruent word-color pairing in a Stroop task (Miyake & Friedman, 2012; Miyake et al., 2000; Stroop, 1935). However, these same tasks are considered by other researches to be related to attention control more



broadly, and as such would be likely captured by complex span measures (Engle & Kane, 2004; Unsworth & Spillers, 2010).

As such, there are instances in which researchers do not find any additional benefit to “inhibition,” but do find advantages of other tasks for which forgetting of outdated information is advantageous, and potentially obscuring an important theoretic relationship. For example, Christopher et al. (2012) found no direct relationship between measures of response inhibition and reading comprehension. However, when a measure of general intelligence, was added to their analyses, general intelligence explained unique variance in reading comprehension beyond working memory capacity. These results could reflect a situation in which variance was obscured by measuring a process along a single task dimension, but when domain-general measures were used the effect is present. We would argue that the variance accounted for by general intelligence suggests that organized forgetting, or disengagement, is present and important for performance, but is overlooked if the results are limited to contributions of complex span tasks and response inhibition measures. This final point brings us to our third latent factor: working memory updating.

### Maintenance, Disengagement, and Updating

Shipstead et al. (2016) define disengagement as a general process of the central executive; however, they did not test this theory by including factors other than fluid intelligence that may also rely on disengagement processes. To examine whether or not disengagement occurs in tasks other than those used to measure fluid intelligence, we elected to include a working memory updating factor that we believe also reflects both maintenance and disengagement, and likely to a more balanced extent than measures of fluid intelligence, which are more heavily oriented toward disengagement.

Memory updating is the process responsible for creating representations of information entering working memory as well as altering the contents of working memory to swap out irrelevant information with information more suited to ongoing cognition (Miyake et al., 2000). This process is strongly related to working memory capacity in that representations must be maintained as they are encountered (Miyake et al., 2000; Oberauer, Süß, Wilhelm, & Sander, 2007). In updating tasks, however, there is an additional need to remove outdated information from memory (active deletion), as well as substitute the target (Ecker et al., 2010). This substitution process is not advantageous in traditional working memory tasks as one must remember all information presented in a series, with no replacement between presentation and recall episodes for that trial. However, updating tasks, such as the running span task or the keeping track task, require replacement of information between the presentation of that information and the recall episode within trials.

Was et al. (2011) found that content-embedded tasks, in which participants had to mentally rearrange word lists in alphabetical order, was a stronger predictor of comprehension than complex span tasks in which participants maintained and recalled information independent from within-trial processing components. In operation span tasks, for example, individuals must recall a series of letters in serial order. These letters are presented after a processing demanding task in which they must evaluate solutions to a math problem. Thus, the to-be-recalled information is independent of

the processing demands of the math problems in so much as the math problems are not needed for recall later. On the other hand, content-embedded tasks require individuals to process and substitute items being held in working memory that were then recalled. Thus, the added predictive value of the content-embedded task lies in the ability of individuals to release no longer relevant information from memory. We suggest that, this content-embedded task is essentially a measure of working memory updating.

Updating improves STM by reducing proactive interference around the attentional target, the same argument that has been made regarding inhibition and reading comprehension (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999) and processing efficiency in reading comprehension (Borella et al., 2010). Subsequently, updating was added to our model to test whether disengagement can be isolated in constructs other than fluid intelligence. The inclusion of updating allows for very specific predictions regarding the role of disengagement in complex cognition. Specifically, if fluid intelligence and updating both capture similar aspects of performance (disengagement), updating should account for the same residual variance in reading comprehension as fluid intelligence. On the other hand, if fluid intelligence captures variance related to another process relevant to reading comprehension, it will maintain predictive value even after the inclusion of updating.

### Maintenance and Disengagement in Other Complex Verbal Abilities

Finally, we examined one additional measure of educationally based language outcomes in this study: second-language vocabulary learning.<sup>1</sup> While there is a long history of research on processes underlying reading comprehension dating back to the 1930s (see Gray, 1941 for a summary of early work on reading comprehension), the same cannot be said for vocabulary learning in adults. Of particular interest to us was foreign-language vocabulary learning. For this measure, we created a vocabulary-learning paradigm where participants were presented 20 Arabic (auditory) English (visual) word pairs, and then an immediate test on 15 of the previously presented pairs. After 90 min had passed, they received a surprise assessment of all 20 word pairs. Performance on this surprise recognition task was used as our dependent measure.

Although this portion of the study was more exploratory in nature, we suspected that the role of working memory capacity and fluid intelligence would be the same for vocabulary learning as reading comprehension. First, recent studies have shown a strong relationship between fluid intelligence and complex associative learning, suggesting that complex associative learning correlates with intelligence as well or better than it does working memory (Tamez, Myerson, & Hale, 2008; Williams & Pearlberg, 2006). Second, assuming fluid intelligence reflects disengagement, this may help isolate the pairing of words with their translations. Keeping word pairs separate would also reduce the retrieval search space (Anderson & Bower, 1974). Finally, Cattell’s Investment

<sup>1</sup> A portion of this project was funded by the Center for Advanced Study of Language (CASL), with the goal of better understanding cognitive processes related to language learning. The vocabulary learning task was included to examine the relationship between cognitive ability and second language vocabulary learning.

Theory (Cattell, 1971, 1987) suggests we would see a relationship between fluid intelligence and crystallized abilities such as vocabulary learning.

Cattell's Investment Theory (Cattell, 1971, 1987) proposes that there is a causal link from fluid intelligence to crystallized intelligence, such that individual differences in crystallized intelligence are related to individual differences in fluid intelligence at the time learning occurs. Arguably, this theory fits reading comprehension, as it is a test of knowledge extracted from a passage, but vocabulary learning is a very straightforward crystallized measure. If greater fluid intelligence does in fact result in greater crystallized intelligence, then we anticipate seeing a stronger predictive relationship from fluid intelligence to performance than from working memory capacity to performance. Additionally, if we can isolate the influence of fluid intelligence independent of working memory capacity, then we can make inferences about the processes that are important for Investment Theory.

### Predictions

We argue that working memory capacity measures have been limited in their ability to explain differences in reading comprehension because of the emphasis on maintenance of information. As such, we anticipated that fluid intelligence would explain reading comprehension above and beyond working memory capacity. We also provided a deeper exploration of the maintenance and disengagement theory of Shipstead et al. (2016) by testing if the "disengagement" variance captured by fluid intelligence measures was process-general in nature. To this end we included a measure of working memory updating to test whether the residual variance accounted for by fluid intelligence was unique to fluid intelligence or common to other theoretical constructs that should rely on disengagement. Finally, we predicted that both maintenance and disengagement as general processes would independently predict performance on reading comprehension and vocabulary learning tasks.

### Method

#### Participants

The sample included 567 people between the ages of 18–35 ( $M = 22.38$ ,  $SD = 4.445$  years) who completed a four-session screening procedure at either Georgia Institute of Technology (Georgia Tech) or Indiana University-Purdue University, Columbus (IUPUC). Of these, 267 were females and 350 were attending college. Participants were fluent in English by age 5 and were not fluent in Arabic. All participants indicated normal hearing and were not taking medications that could affect their attention or memory. Participants were checked for normal or corrected-to-normal vision. For more information about this sample, please refer to Shipstead et al. (2016).

All participants were tested on individual computers in groups of 1–5 as part of an extensive screening procedure. The screening procedure collected data for multiple studies, and not all tasks were analyzed in the present study. This study was reviewed by the Georgia Institute of Technology's Institutional Review Board and was approved under protocol #H12234 entitled: The Relationships

Among Working Memory Tasks And Their Relations To Fluid Intelligence And Higher-Order Cognition.

### Tasks

#### Working memory capacity.

**Automated operation span (Unsworth, Heitz, Schrock, & Engle, 2005).** Test takers recalled a series of presented items, the presentation of which is interrupted by a simple processing task. For the operation span the to-be-remembered items were letters from the English alphabet. The processing task was a simple mathematical equation that must be solved before the next letter of a sequence is presented. Lists lengths varied between three and seven items. The list lengths were presented in a randomized order, with the constraint that a given length could not repeat until all lengths had been presented. Each list length was used three times. The dependent variable was the number of letters recalled in proper serial position during the session (i.e., partial scoring method).

**Automated symmetry span (Unsworth, Redick, Heitz, Broadway, & Engle, 2009).** The to-be-remembered items were spatial locations in  $4 \times 4$  grid. The processing task required test-takers to judge whether or not a figure in an  $8 \times 8$  grid was symmetrical. List lengths were 2–5 items. Other characteristics mirrored the operation span.

**Automated rotation span (Harrison et al., 2013).** The to-be-remembered items were a sequence of long and short arrows, radiating from a central point. The processing task required test-takers to judge whether a rotated letter was forward facing, or mirror-reversed. List lengths were 2–5 items. Other characteristics mirrored the operation span.

**Fluid intelligence.** For all fluid intelligence tasks, the dependent variable was the number of correct responses provided within the allotted time.

**Raven's advanced progressive matrices (Raven, 1990; odd problems).** On each trial, eight abstract figures were embedded in a  $3 \times 3$  matrix. The final position in the matrix was blank. Test-takers selected one of eight options to complete the sequence. Ten minutes were given to solve 18 problems.

**Letter sets (Ekstrom, French, Harman, & Dermen, 1976).** On each trial, five 4-letter strings were presented. Four of the sets followed a specific rule. The test-taker needed to discern this rule and decide which string did not follow it. Seven minutes were given to complete 30 problems.

**Number series (Thurstone, 1938).** A series of numbers was presented on a computer screen. A rule joined these numbers. The test-taker needed to discern this rule and decide which number was next in the sequence. Five minutes were given to complete 15 problems.

#### Memory updating.

**Running span.** A running spatial and two running digit spans were administered. In the running spatial span, test takers saw a  $4 \times 4$  grid. Locations on the grid were highlighted one at a time. Test-takers needed to remember the last 3–7 grid locations that were highlighted, after being instructed before the trial how many of the last  $n$  items they would need to recall for the upcoming series. The dependent variable was the number of locations that were correctly recalled in their proper serial position. The running digit spans included two or four digits that were used for recall

rather than spaces. A  $z$ -score composite was created using the spatial, the two, and the four-digit strings and was used as the indicator for the updating factor to avoid arbitrarily selecting a single stimulus set.

**N-back.** The three n-back tasks, respectively, presented a series of words, faces, or windings in order on a computer monitor. In each task there were 10 targets and 10 lures for each position, and 10 filler items never repeated. The test-taker judged whether the currently presented stimuli matched the stimulus that was presented three items ago. The task also included a 40 trial practice. Each stimulus was presented for up to 2,000 ms, with a 500 ms pause. The dependent variables were hits, false alarms, and  $d'$  that was calculated using hits at the three-back position, and false alarms at each of the lure positions (2, 4, and 5). A  $z$ -score composite was created based on all three n-back scores and was used as the n-back indicator for the updating factor to avoid arbitrarily selecting a single stimulus set.

**Keeping track (Yntema & Trask, 1963).** This task contained six categories (countries, relatives, metals, animals, colors, and distances). At the start of each trial, 2–6 of these categories (e.g., metals, animals, and countries) were displayed for the test-taker to keep track of, followed by a list of words from all six categories (e.g., iron, cow, England). Test takers were required to remember the most recent word from each of the categories chosen for a particular trial. Each category was tested three times. The dependent variable was the total number of correct responses across 15 trials.

#### Outcome measures.

**Reading comprehension.** In the reading comprehension task, test-takers read a paragraph (~140 words) that told a short story. After reading the paragraph, the test-takers advanced to three questions via mouse click. The paragraph was no longer visible when the questions were answered. The first question referred to a pronoun that appeared in the last sentence of the paragraph (e.g., “He paid his bill and left”; “who was ‘he’?”). This pronoun, in turn, referenced a noun that had appeared 4–7 sentences before the pronoun. The second question was a fact-based question that referenced an event somewhere in the paragraph. Both of these questions were answered by clicking on one of four responses. The third question was a true/false question regarding the events in the paragraph. The dependent variable was accuracy on each question over eight paragraphs.

**Vocabulary learning.** Test-takers learned an association between 20 common English-Arabic concrete nouns during an initial learning phase. After a delay of about 90 min, a surprise recognition test was administered to assess the number of English-Arabic pairs that were maintained in memory without active practice.

During the initial learning phase, test-takers completed four alternating study and test blocks. Each study block contained 20 trials (i.e., English-Arabic word pairs) in which the same set of English and Arabic words were presented. English word was displayed in white text in the center of the screen. After 500 ms, the Arabic equivalent was spoken via headset while the English word was still displayed on the screen. In total, the English word remained visible for 2,500 ms before the screen cleared. The display remained blank for 1,000 ms before a fixation screen appeared informing test-takers that a new word was about to be presented. Pairs were randomly presented throughout each study block. Arabic words were recorded using Google Translate where

Abjad characters were entered into the search engine and the Modern Standard Arabic translation was chosen. All Arabic words were spoken using a male voice. Irrelevant sounds that arose from the recording software (e.g., electronic pops and clicks) were edited out.

Immediately after the study block was a test block in which 15 of the original 20 pairs were tested. One of the 15 Arabic words was spoken via headset without the English word on the screen. Test-takers were informed that they could listen to the Arabic word as many times as they needed by clicking the screen. Once test-takers were ready, they could click a box labeled “Done” to advance to a recognition screen. The recognition screen contained 20 boxes, one box for each English word that was presented during study. Test-takers were instructed to click on the box that contained the English translation of the Arabic word they just listened to. Feedback regarding accuracy was not provided. Once a selection was made, the process would repeat until all of the 15 selected pairs had been tested. While presented in random order, the same 15 Arabic-English pairs were tested for all participants and are referred to as “tested pairs.” Thus, five Arabic-English pairs were never tested during the initial learning phase and are referred to as “untested pairs.” After this, the study block would repeat followed by the test block until there were four presentations of both study and test blocks.

Participants continued to complete cognitive tasks for roughly 90 min after the initial learning phase. After this time had elapsed, a surprise recognition test was administered in which all 20 English-Arabic pairs were tested. Test-takers were not informed of this test nor were they told to practice the English-Arabic pairs they had learned during the initial learning phase. The format of this test was identical to the test block of the initial learning phase. Neither a preceding study period nor feedback regarding accuracy was provided. The dependent variable was the total number of accurately identified English-Arabic pairs to make a total of 20 points possible (5 points for untested pairs, 15 points for tested pairs).

Structural equation modeling was used to investigate the processes (i.e., working memory capacity, fluid intelligence, and working memory updating) that were the best predictors of language comprehension and second-language vocabulary learning and performance. Models incorporating both working memory capacity as well as fluid intelligence were tested, to assess whether models including updating as well as fluid intelligence provided a more comprehensive explanation of language proficiency than those using working memory capacity as the mediating variable.

**Data prescreening and preparation.** Univariate outliers were identified as any individual scores that exceeded the grand mean by more than 3.5  $SD$ s. Out of more than 18,000 individual observations, only eight observations fit this criterion. The scores for these data were replaced using the cutoff value of  $\pm 3.5 SD$ .

Random events such as equipment malfunction and experimenter error occasionally resulted in data loss for a given participant. Missing scores were imputed using the maximum likelihood function in EQS 23 (under 1% of all observations). Multivariate analyses were conducted using EQS.

**Reported fit statistics.** Multiple fit indices are reported for each model;  $\chi^2/df$  served as our “badness-of-fit” statistic. This statistic cannot be formally interpreted because it is subject to sample size, but we accepted values up to 3 given our large sample size (Kline, 1998). Root mean square error of approximation



(RMSEA) estimates the model fit to the population. Standardized root mean square residual (SRMR) reflects average deviation of the reproduced covariance matrix from the observed matrix. For both RMSEA and SRMR, values below .05 are ideal, but up to .08 are acceptable (Browne & Cudeck, 1993; Kline, 1998). Non-normed fit index (NNFI) and comparative fit index (CFI) make a comparison of the hypothesized model to a model in which observed variables are assumed to be uncorrelated. Values above .95 represent a good fit (Hu & Bentler, 1999).

## Results

Descriptive statistics, factor analysis, and latent factor correlations are presented in the following tables. Table 1 shows descriptive statistics for all tasks used to create the latent variables of working memory capacity, fluid intelligence, updating, reading comprehension, and second-language vocabulary learning. The wide range reflects our efforts to recruit individuals outside of a university setting.

Confirmatory factor analysis was used to verify that tasks loaded onto the factors of working memory capacity, fluid intelligence, updating, reading comprehension, and vocabulary learning as we anticipated. Task loadings for each factor are shown in Table 2. All loadings were high for each factor, and indicated robust, coherent factors.

Correlations among latent factors are included in Table 3, and in structural equation models to follow as appropriate. As anticipated correlations between working memory capacity, updating, and fluid intelligence were high, as were the relationships between these cognitive factors and our reading comprehension and vocabulary learning factors.

Table 1  
*Descriptive Statistics*

Task	<i>M</i>	<i>SD</i>	Range	Skew	Kurtosis
Ospan	54.23	5.52	0–75	–.89	.22
SymSpan	26.72	9.02	3–42	–.43	–.48
RotSpan	24.70	9.72	0–42	–.43	–.59
RunDig2	55.40	20.83	0–108	–.43	–.07
RunDig4	56.83	19.27	0–107	–.49	.32
SpatialRun	24.61	11.26	0–49	–.32	–.78
NBackWing	.74	1.68	–4.54–6.69	.45	1.91
NBackWord	1.07	1.63	–4.37–6.89	.61	2.30
NBackFace	.33	1.43	–4.60–5.38	–.37	1.55
KeepTr	33.03	10.91	4–54	–.46	–.69
Raven	8.70	3.91	0–18	–.07	–.88
LetterSet	15.43	5.26	1–29	–.03	–.60
NumSeries	8.57	3.58	0–15	–.22	–.86
Pronom	4.99	2.02	.00–8.00	–.47	–.59
Fact	5.57	1.98	.00–8.00	–.71	–.39
T/F	5.93	1.26	2.00–8.00	–.66	.16
Untested Pairs	3.73	5.54	0–5	.74	1.38
Tested Pairs	10.91	6.27	9–20	–.43	–.81

*Note.* Ospan = operation span; SymSpan = symmetry span; RotSpan = rotation span; Pronom = pronominal reference component of reading comprehension; Fact = accuracy for questions on the reading comprehension; T/F = accuracy on the true or false questions for reading comprehension; Tested Pairs = accuracy for vocabulary words that were tested during learning; Untested Pairs = accuracy for untested vocabulary words that were not tested at learning. Mean, *SD* and range were calculated before the removal of outliers.

Table 2  
*Confirmatory Factor Analysis Task Loadings on Each Factor*

Task	WMC	Update	Gf	Vocab	Reading
Ospan	.69				
SymSpan	.75				
RotSpan	.78				
Nback		.79			
RunningSpan		.59			
Keeping Track		.79			
Ravens			.81		
Letter Series			.79		
Number Series			.82		
Tested Pairs				.95	
Untested Pairs				.85	
Pronom					.84
Fact					.85
T/F					.60

*Note.* WMC = working memory capacity; Update = updating; Gf = fluid intelligence; Vocab = vocabulary learning; Reading = paragraph comprehension. For row headings: Ospan = operation span; SymSpan = symmetry span; RotSpan = rotation span; Tested Pairs = accuracy for vocabulary words that were tested during learning; Untested Pairs = accuracy for vocabulary words on that were not tested during learning; Pronom = pronominal reference component of reading comprehension; Fact = accuracy for questions on the paragraph comprehension; T/F = accuracy on true false questions for reading comprehension.

## Structural Equation Models

Table 4 shows fit statistics for all to-be-reported models. The fit for all models was considered acceptable.

## Contributions of Working Memory Capacity and Fluid Intelligence to Reading Comprehension

Our first analysis examined the individual correlations of working memory capacity and fluid intelligence to reading comprehension. As can be seen in Figure 4, fluid intelligence (Gf) had a strong relation to reading comprehension (.62), while working memory capacity had no direct correlation beyond its substantial association with fluid intelligence. This is not to say that working memory capacity and reading comprehension are independent. Instead, it indicated that their correlation is explained by variance that working memory capacity shared with fluid intelligence.

In Figure 5, we examined independent contributions of processes related to both complex span measures of working memory capacity and fluid intelligence measures. In this model we cross-

Table 3  
*Confirmatory Factor Analysis: Correlations Amongst Latent Factors*

Task	WMC	Update	Reading	Vocab	Gf
WMC	1				
Update	.86	1			
Reading	.59	.77	1		
Vocab	.68	.61	.48	1	
Gf	.85	.90	.63	.83	1

*Note.* WMC = working memory capacity; Update = memory updating; Reading = reading comprehension; Vocab = vocabulary learning; Gf = fluid intelligence.

Table 4  
Model Fit Statistics

Model	$\chi^2$	df	$\chi^2/df$	RMSEA	SRMR	NNFI	CFI
WMC Gf and Reading Comp	65.05	24	1.71	.06	.06	.93	.98
Maintenance, disengagement and Reading Comp	64.10	22	1.51	.06	.05	.97	.98
WMC, Gf, Updating, and Reading Comp	136.70	44	3.49	.06	.05	.97	.97
WMC, Gf, and Vocab Learning	111.63	24	1.71	.07	.05	.97	.98
Maintenance, disengagement and Vocab Learning	56.07	15	3.21	.07	.06	.92	.96
WMC, Gf, Updating, and Vocab Learning	100.31	34	3.49	.06	.03	.94	.98

Note. CFI = comparative fit index; RMSEA = root mean square error or approximation; SRMR = standardized root mean square residual; NNFI = nonnormed fit index; WMC = working memory capacity; Update = memory updating; Reading = reading comprehension; Vocab = vocabulary learning; Gf = fluid intelligence.

loaded all fluid intelligence tasks onto the working memory capacity factor. Thus, the factor labeled “Gf-res” reflected residual variance in the fluid intelligence factor that was “independent” of working memory capacity. The result allowed us to see that the shared maintenance components of working memory capacity and fluid intelligence measures explained about 29% of the variance in reading comprehension (obtained by squaring the regression path or  $(.54)^2$ ). Fluid intelligence residual or disengagement added another 8% above-and-beyond working memory capacity.

Our prediction was that this added 8% variance was explained by general disengagement processes. To test that hypothesis, our final model added an updating factor, based on the need to both maintain information and disengage from information in these tasks. Further, the updating tasks were more similar in nature to the complex span measures of working memory capacity. Therefore, we predicted the updating factor would provide a clearer dissociation between maintenance and disengagement. If residual fluid intelligence variance was still significant, then specific processes related to those tasks predicted reading comprehension. However, if residual fluid intelligence variance was no longer predictive of performance, then we could infer that the updating residual was now reflecting the same process general disengagement variance.

As can be seen in Figure 6, the updating tasks were cross-loaded onto working memory capacity. This was done because updating

tasks tend to have a strong working memory component (Miyake et al., 2000) as well as a disengagement component via active substitution (Ecker et al., 2010). It was, therefore, necessary to isolate this disengagement process from these tasks to refine the factor such that it primarily reflected the act of updating memory, and not maintaining information.

In line with our prediction, updating not only had a unique correlation to reading comprehension, but this relationship also accounted for any direct relationship between fluid intelligence and reading comprehension. In other words, in terms of its relationship to reading comprehension, fluid intelligence represented a mix of domain-general maintenance processes (as reflected by working memory capacity) and general disengagement processes.

### Working Memory Capacity and Fluid Intelligence in Vocabulary Learning

The previous analyses were duplicated for the vocabulary learning task. All model fit was considered “good” with a CFI of .98, and an RMSEA lower than .08.

Our first model for vocabulary learning examined the individual correlations of working memory capacity and fluid intelligence to performance on a surprise vocabulary quiz, approximately 90 min after word learning. Based on Cattell’s Investment Theory, we

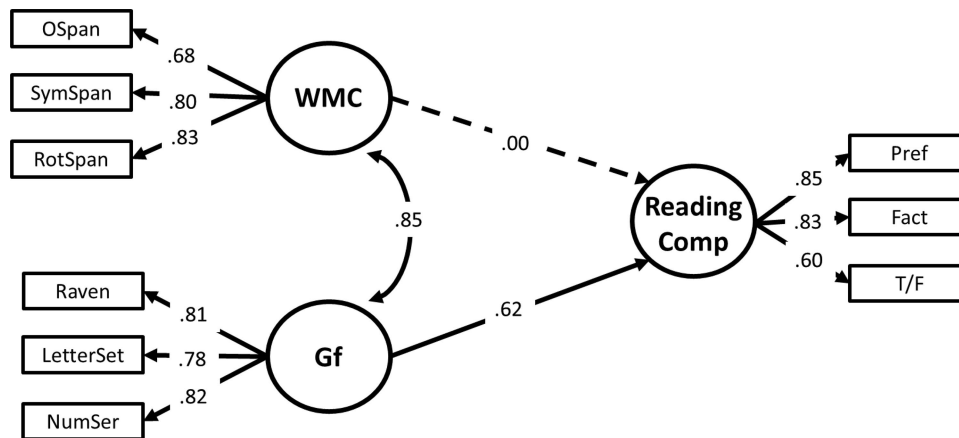


Figure 4. Fit statistics: Model fit was good:  $\chi^2 = 65.05$  (24);  $p < .001$ ; comparative fit index (CFI) = .98; root mean square error or approximation (RMSEA) = .05; standardized root mean square residual (SRMR) = .05; nonnormed fit index (NNFI) = .98. Only the path from fluid intelligence to reading comprehension performance was significant. This path accounted for almost 38% of the variance in reading comprehension.

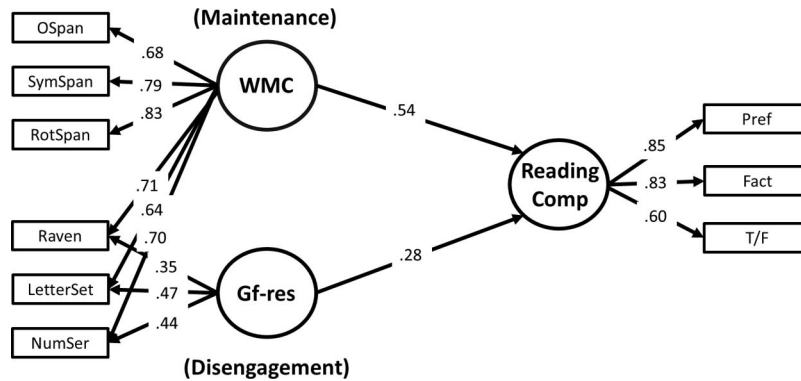


Figure 5. Model fit was good. Fit statistics:  $\chi^2 = 64.10$  (22);  $p < .001$ ; comparative fit index (CFI) = .98; root mean square error or approximation (RMSEA) = .06; standardized root mean square residual (SRMR) = .05; nonnormed fit index (NNFI) = .97. Both paths were significant with “maintenance” related variance (i.e., the working memory capacity factor with the additional fluid intelligence measures) accounting for 29% of the variance in reading comprehension performance. The Gf residual or “disengagement” related variance (i.e., variance in the three fluid intelligence measures that was not shared with the complex span measures) accounted for an additional 7.8% of the variance in reading comprehension performance.

anticipated fluid intelligence would predict performance beyond working memory capacity. As can be seen in Figure 7, fluid intelligence had a strong relation to vocabulary learning (.77), while working memory capacity had no direct correlation. As with reading comprehension, this relationship indicated that the relationship between working memory capacity and vocabulary learning was explained by variance shared with fluid intelligence.

### Maintenance and Disengagement in Vocabulary Learning

As can be seen in Figure 8, the fluid intelligence tasks were again cross-loaded onto working memory capacity that resulted in a shared maintenance factor and a fluid intelligence residual or disengagement factor. Once again, when factors are set up to examine contributions at the process level, both maintenance and

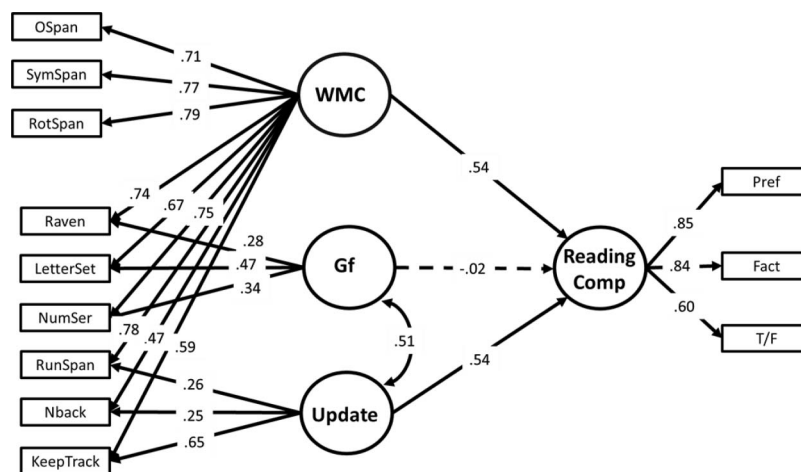


Figure 6. Fit statistics:  $\chi^2 = 136.70$  (44);  $p < .001$ ; comparative fit index (CFI) = .97; root mean square error or approximation (RMSEA) = .06; standardized root mean square residual (SRMR) = .05; nonnormed fit index (NNFI) = .97. Both paths were significant with “maintenance” related variance (i.e., the working memory capacity factor with the additional fluid intelligence measures) accounting for 29% of the variance in reading comprehension performance. The Update residual or “disengagement” related variance (i.e., variance in the three updating measures that was not shared with the complex span measures) accounted for an additional 29% of the variance in reading comprehension performance. The residual Gf was no longer significant. Subsequently, the variance previously accounted for by the fluid intelligence residual is now reflected by the updating residual factor.

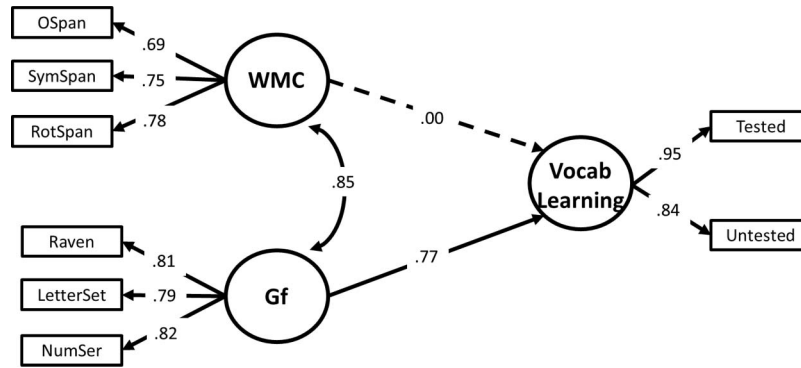


Figure 7. Fit statistics:  $\chi^2 = 111.63$  (24);  $p < .001$ ; comparative fit index (CFI) = .98; root mean square error or approximation (RMSEA) = .07; standardized root mean square residual (SRMR) = .05; nonnormed fit index (NNFI) = .97. Only the path from fluid intelligence to vocabulary learning performance was significant. This path accounted for 59.3% of the variance in vocabulary learning.

disengagement accounted for significant, independent variance in vocabulary learning performance.

Additionally, as can be seen in Figure 9, the updating tasks were again added and also cross-loaded onto working memory capacity, or “maintenance.” In line with our prediction, updating not only showed a unique relationship to vocabulary learning, but this relationship also accounted for any correlation between fluid intelligence and reading comprehension. In other words, fluid intelligence represented a mix of maintenance processes (as reflected by working memory capacity) and disengagement processes (as reflected by updating).

### Summary and Discussion

This tested a process general approach to explaining individual differences in reading comprehension and second language vocabulary learning. Complex span measures of working memory capacity, measures of fluid intelligence, and working memory updating were included based on a novel framework of maintenance

and disengagement proposed by Shipstead et al. (2016). Our results not only expanded this framework, but also tested its ability to provide an explanation of reading comprehension and vocabulary learning. First, we showed that fluid intelligence was a better predictor of performance than working memory capacity when the two constructs were compared directly. Second, we demonstrated that the predictive contribution of fluid intelligence and working memory capacity, in terms of reflecting both maintenance and disengagement, were both separable through the use of structural equation modeling, and significant in predicting independent variance in performance. Third, we examined the degree to which disengagement proposed by Shipstead et al. (2016) is also a process relevant to theoretical constructs other fluid intelligence (i.e., updating). Our results indicated that disengagement is a general process independent of fluid intelligence measures and that using measures that tap this process increase the prediction of performance beyond measures of working memory capacity alone.

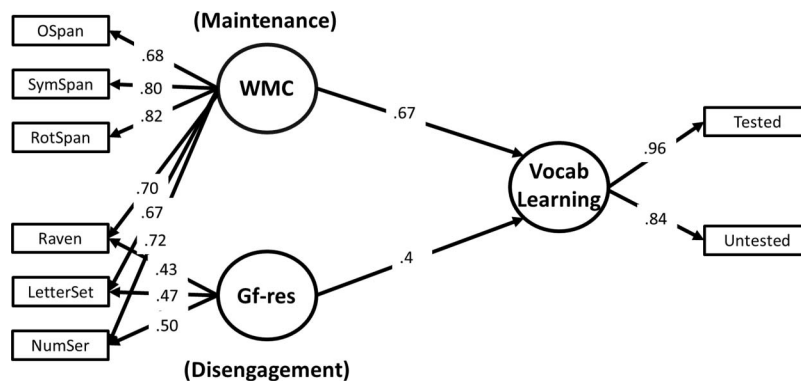


Figure 8. Model fit was good. Fit statistics:  $\chi^2 = 56.07$  (15);  $p < .001$ ; comparative fit index (CFI) = .98; root mean square error or approximation (RMSEA) = .07; standardized root mean square residual (SRMR) = .05; nonnormed fit index (NNFI) = .97. Both paths were significant with “maintenance” related variance (i.e., the working memory capacity factor with the additional Gf measures) accounting for 44.9% of the variance in vocabulary learning performance. The Gf residual or “disengagement” related variance (i.e., variance in the three fluid intelligence measures that was not shared with the complex span measures) accounted for an additional 16% of the variance vocabulary learning performance.



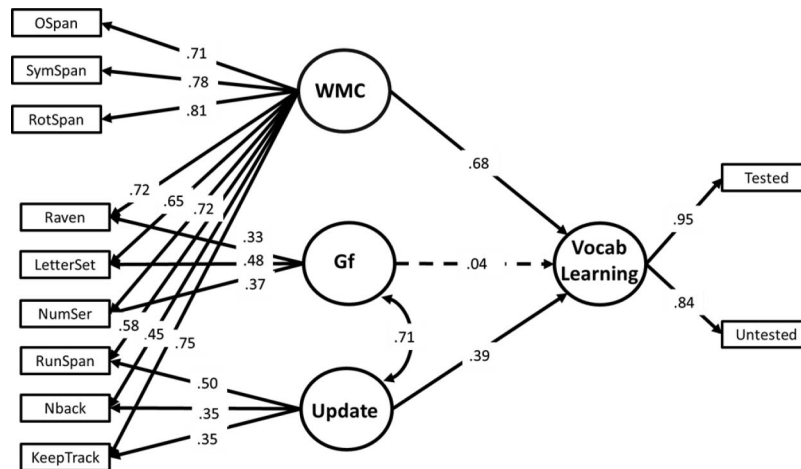


Figure 9. Fit statistics:  $\chi^2 = 100.31$  (34);  $p < .001$ ; comparative fit index (CFI) = .98; root mean square error or approximation (RMSEA) = .07; standardized root mean square residual (SRMR) = .03; nonnormed fit index (NNFI) = .97. Both paths were significant with “maintenance” related variance (i.e., the working memory capacity factor with the additional fluid intelligence measures) accounting for 46.2% of the variance in vocabulary performance. The Update residual or “disengagement” related variance (i.e., variance in the three updating measures that was not shared with the complex span measures) accounted for an additional 15.2% of the variance in reading comprehension performance. The residual Gf was no longer significant. Subsequently, the variance previously accounted for by the fluid intelligence residual is now reflected by the updating residual factor.

Whereas Daneman and Merikle (1996) argued that working memory capacity tasks predict language comprehension because complex span measures engage executive functions that maintain information when distraction is high, we interpreted our results to mean that this only reflects half of the story. Our results showed that “disengagement” also aided in performance by preventing reactivation of now uninformative stimuli. Moreover, as proposed by Shipstead et al. (2016), disengagement is a process that complex span measures do not capture well. Subsequently, including tasks which capture disengagement processes (measured via fluid intelligence or working memory updating tasks) provides a robust explanation of individual differences in reading comprehension. The role of disengagement and its link to reading comprehension and verbal fluency also complements findings of Was and colleagues (2007), which showed that available long-term memory predicted reading comprehension beyond both working memory span tasks as well as background knowledge. Specifically, some of the measures of available long-term memory used by Was et al. were similar to verbal fluency tasks as they measure retrieval of previously presented categorical exemplars. As such, these available long-term memory measures may also tap into disengagement processes, as perseveration or re-retrieval would result in poorer performance on these measures as well. This additional need for disengagement may explain why available long-term memory measures predicted reading comprehension better than working memory capacity. Further studies should investigate this relationship in more detail.

Given the theoretical foundation of Shipstead et al. (2016), our results suggest that we should be examining more fluid measures that incorporate disengagement functions when examining complex phenomenon such as reading comprehension ability. Further, these findings corroborate work by Was and colleagues (2011)

suggesting that updating may also be a better indicator of performance on reading comprehension measures than complex span measures, and are also consistent with a study by Chen and Li (2007) who showed that updating, but not processing speed predicted unique variance in fluid intelligence. Specifically, measures of working memory updating emphasize active deletion of outdated material, as well as substitution of old material with new, more relevant information, rather than just enhanced speed of processing, or a set memory load.

By examining our theoretical constructs at the latent level, rather than task level, we were also better able to measure broad theoretical constructs in a reliable way. What we mean by this is that many measures of processing ability of inhibition rely on reaction time (RT) difference scores. While it is beyond the scope of this article to go into detail on this matter, the use of accuracy-based tasks in construct creation is much more reliable and consistent in findings across populations (see Draheim, Mashburn, Martin, & Engle, in press). The use of constructs such as maintenance and disengagement to describe the way individuals manage a complex series of information is advantageous for other reasons as well. First, some individuals may approach disengagement from different mechanistic standpoints and make use of different strategies vis-à-vis disengagement. As such, using very specific indicators may not provide the best evidence regarding general methods of performance across diverse populations. Second, our incorporation of updating suggests that disengagement is a general process that is not specific to a single task set, and as such cannot be solely attributed to task specific variance on measures of fluid intelligence.

One of the major contributions of this study is the extension of the maintenance and disengagement framework to constructs beyond those from which it was originally designed. Specifically, we

were able to isolate variance contributed by maintenance and disengagement through modeling techniques. Then we tested whether or not residual variance from fluid intelligence measures was independent of that construct. In doing so, we showed that disengagement is a general process independent of fluid intelligence. Further, we can argue with relative certainty that, in this context, fluid intelligence does not provide any informative prediction beyond this general process of disengagement. This finding is instrumental in supporting the cohesiveness of this disengagement process across task types and other theoretical constructs such as updating.

The second major contribution of this study is the use of the maintenance and disengagement framework of Shipstead et al. (2016) to predict real-world performance. Not only were we able to use these constructs to independently predict reading comprehension, but also second-language vocabulary learning. We anticipated a similar predictive relationship of working memory capacity and fluid intelligence to that found in reading comprehension, because of their relatively concrete task nature. As anticipated, fluid intelligence predicted performance beyond working memory capacity. Additionally, this predictive advantage was because of variance common to both updating and fluid intelligence measures, which we call disengagement. Finally, both "maintenance" and "disengagement" were significant predictors of performance. These parallel results are important for several reasons. First, they further support the idea that fluid intelligence measures can be used to capture variance that is separable from working memory capacity (Shipstead et al., 2016). Second, our results show that this residual variance captured by both fluid intelligence and updating measures (disengagement) is a general process that impacts performance on multiple kinds of verbal tasks. Third, these data provide support for Cattell's Investment Theory.

Cattell's Investment Theory (Cattell, 1987) proposed that fluid intelligence at the time the learning occurred accounts for differences in crystallized intelligence measured later. Crystallized intelligence is best measured by performance on our vocabulary learning task. Based on this theory, we anticipated a strong relationship between fluid intelligence and vocabulary learning after a delay. In line with Cattell's Investment Theory, and our predictions, our measure of vocabulary learning was strongly predicted by fluid intelligence. Subsequently, we can infer that disengagement is relevant to Investment Theory as well.

The specific role of disengagement in vocabulary learning is not addressed in this article; however, we can speculate on these functions. For example, when learning word pairings, one needs to not only remember the new association, but also to avoid distraction or competition from previously presented word pairs and associations with similar words. Disengagement is in turn congruent with explanations for the phenomenon attributed to inhibition by Hasher and Zacks (1988) such as separating competing representations (i.e., maintaining other word pairs separately, and in turn reducing proactive interference), reducing intrusion errors, and reducing the search space for previously maintained information (see Shipstead et al. [2016] for additional discussion regarding the functions of disengagement).

In summary, our results showed that the general processes of maintenance and disengagement can be separated using structural equation modeling techniques. Second, we showed that a large percentage of the variance underlying reading comprehension performance can be explained in terms of the maintenance and dis-

engagement framework and processes proposed by Shipstead et al. (2016). Although this is not the only or a complete explanation of the processes underlying reading comprehension, we believe this to be a very theoretically sound, reliable, and parsimonious explanation for a large amount of variance in performance particularly when a diverse population sample is used. Our results are parsimonious in that they approach the underlying functions of reading comprehension from a broad-process perspective rather than a deconstructive mechanistic perspective. This allows for an in-depth account of performance to the extent that, individuals may use a variety of mechanisms toward the end of maintaining or disengaging from information. As such, the process-general approach allows for a more comprehensive evaluation of the processes underlying performance across a wide range of abilities.

Moreover, our findings extended our understanding of the types of latent constructs which reflect maintenance and disengagement, beyond measures of working memory capacity and fluid intelligence. For example, had the predictive value of fluid intelligence remained significant following the addition of updating, then we would have concluded that other aspects of performance measured by fluid intelligence were important for predicting reading comprehension. Further, we can now make some more educated guesses regarding the mechanisms that are important for reading comprehension, and also shared between measures of fluid intelligence and working memory updating, but which were not common to complex span measures.

Finally, the extension of this theoretical framework and its ability to predict ability to learn second-language vocabulary items further extends the validity of the maintenance and disengagement paradigm to explaining performance in real-world educational scenarios. Although our results were robust for the constructs and outcomes measured here, the degree to which these findings extend to other reading comprehension measures, and other language learning processes needs further investigation. For example, it cannot be directly inferred that these results will translate to other reading comprehension measures such as garden path tasks (MacDonald, Pearlmutter, & Seidenberg, 1994), or more complex language abilities such as grammatical proficiency.

Future directions should include a clarification of the processes disengagement captures, as well as contexts in which currently held beliefs about working memory capacity advantages based on complex-span measures are missing the influence of disengagement processes. Additionally, we are not arguing that this framework provides a complete and comprehensive framework for understanding the processes that underlie reading comprehension performance. However, we do suggest that a more process general approach may be advantageous in context in which a wide range of abilities are included.

## References

- Anderson, J. R., & Bower, G. H. (1974). A propositional theory of recognition memory. *Memory & Cognition*, 2, 406–412. <http://dx.doi.org/10.3758/BF03196896>
- Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 119.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47–89. [http://dx.doi.org/10.1016/S0079-7421\(08\)60452-1](http://dx.doi.org/10.1016/S0079-7421(08)60452-1)

- Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities, 43*, 541–552. <http://dx.doi.org/10.1177/0022219410371676>
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Budd, D., Whitney, P., & Turley, K. J. (1995). Individual differences in working memory strategies for reading expository text. *Memory & Cognition, 23*, 735–748. <http://dx.doi.org/10.3758/BF03200926>
- Cattell, R. B. (1943). The description of personality. I. Foundations of trait measurement. *Psychological Review, 50*, 559–594. <http://dx.doi.org/10.1037/h0057276>
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Oxford, England: Houghton Mifflin.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam, the Netherlands: Elsevier.
- Chen, T., & Li, D. (2007). The roles of working memory updating and processing speed in mediating age-related differences in fluid intelligence. *Neuropsychology, Development, and Cognition, 14*, 631–646. <http://dx.doi.org/10.1080/13825580600987660>
- Chiappe, P., Siegel, L. S., & Hasher, L. (2000). Working memory, inhibitory control, and reading disability. *Memory & Cognition, 28*, 8–17. <http://dx.doi.org/10.3758/BF03211570>
- Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., . . . Olson, R. K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology: General, 141*, 470–488. <http://dx.doi.org/10.1037/a0027375>
- Daneman, M., & Carpenter, P. A. (1980). Individual difference in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450–466. [http://dx.doi.org/10.1016/S0022-5371\(80\)90312-6](http://dx.doi.org/10.1016/S0022-5371(80)90312-6)
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433. <http://dx.doi.org/10.3758/BF03214546>
- Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2018). *Reaction time in differential and developmental research: A review and commentary on the problems and alternatives*. Manuscript submitted for publication.
- Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language, 74*, 77–90. <http://dx.doi.org/10.1016/j.jml.2013.09.003>
- Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 170–189. <http://dx.doi.org/10.1037/a0017891>
- Ekstrom, R. B., French, J. W., Harman, M. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19–23. <http://dx.doi.org/10.1111/1467-8721.00160>
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972–992. <http://dx.doi.org/10.1037/0278-7393.18.5.972>
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–199). New York, NY: Elsevier.
- 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. <http://dx.doi.org/10.1037/0096-3445.128.3.309>
- Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186–204. <http://dx.doi.org/10.1016/j.cortex.2016.04.023>
- Gray, W. (1941). Summary of Reading Investigations: July 1, 1939 to June 30, 1940. *The Journal of Educational Research, 34*, 401–443. <http://dx.doi.org/10.1080/00220671.1941.10881015>
- Harrison, T. L., Shipstead, Z., & Engle, R. W. (2015). Why is working memory capacity related to matrix reasoning tasks? *Memory & Cognition, 43*, 389–396.
- 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.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York, NY: Academic Press.
- Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriath (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge, MA: The MIT Press.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130*, 169–183. <http://dx.doi.org/10.1037/0096-3445.130.2.169>
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 66–71. <http://dx.doi.org/10.1037/0033-2909.131.1.66>
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
- MacDonald, M. C., Pearlmuter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution [corrected]. *Psychological Review, 101*, 676–703. <http://dx.doi.org/10.1037/0033-295X.101.4.676>
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*, 8–14. <http://dx.doi.org/10.1177/0963721411429458>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100. <http://dx.doi.org/10.1006/cogp.1999.0734>
- Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General, 134*, 368–387. <http://dx.doi.org/10.1037/0096-3445.134.3.368>
- Oberauer, K., & Lewandowsky, S. (2016). Control of information in working memory: Encoding and removal of distractors in the complex-span paradigm. *Cognition, 156*(Suppl. C), 106–128. <http://dx.doi.org/10.1016/j.cognition.2016.08.007>
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex



- span. *Psychonomic Bulletin & Review*, 19, 779–819. <http://dx.doi.org/10.3758/s13423-012-0272-4>
- Oberauer, K., Süß, H. M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 49–75). Cambridge, UK: Oxford University Press.
- Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.
- Raaijmakers, J. G., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. *Psychology of Learning and Motivation*, 14, 207–262. [http://dx.doi.org/10.1016/S0079-7421\(08\)60162-0](http://dx.doi.org/10.1016/S0079-7421(08)60162-0)
- Raven, J. C. (1990). *Advanced progressive matrices*. Oxford: Oxford Psychological Press.
- Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 501–526. <http://dx.doi.org/10.1037/xlm0000450>
- Rizzo, N. D. (1939). Studies in visual and auditory memory span with specific reference to reading disability. *Journal of Experimental Education*, 8, 208–244. <http://dx.doi.org/10.1080/00220973.1939.11010163>
- Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, 126, 211–227. <http://dx.doi.org/10.1037/0096-3445.126.3.211>
- Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, 11, 771–799. <http://dx.doi.org/10.1177/1745691616650647>
- Shipstead, Z., Redick, T. S., Hicks, K. L., & Engle, R. W. (2012). The scope and control of attention as separate aspects of working memory. *Memory*, 20, 608–628. <http://dx.doi.org/10.1080/09658211.2012.691519>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. <http://dx.doi.org/10.1037/h0054651>
- Tamez, E., Myerson, J., & Hale, S. (2008). Learning, working memory, and intelligence revisited. *Behavioural Processes*, 78, 240–245. <http://dx.doi.org/10.1016/j.beproc.2008.01.008>
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127–154. [http://dx.doi.org/10.1016/0749-596X\(89\)90040-5](http://dx.doi.org/10.1016/0749-596X(89)90040-5)
- 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. <http://dx.doi.org/10.1016/j.jml.2005.06.003>
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498–505. <http://dx.doi.org/10.3758/BF03192720>
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17, 635–654. <http://dx.doi.org/10.1080/09658210902998047>
- Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, 62, 392–406. <http://dx.doi.org/10.1016/j.jml.2010.02.001>
- Verguts, T., & de Boeck, P. (2002). The induction of solution rules in Raven's Progressive Matrices Test. *European Journal of Cognitive Psychology*, 14, 521–547. <http://dx.doi.org/10.1080/09541440143000230>
- Was, C. A., Rawson, K. A., Bailey, H., & Dunlosky, J. (2011). Content-embedded tasks beat complex span for predicting comprehension. *Behavior Research Methods*, 43, 910–915. <http://dx.doi.org/10.3758/s13428-011-0112-x>
- Was, C. A., & Woltz, D. J. (2007). Reexamining the relationship between working memory and comprehension: The role of available long-term memory. *Journal of Memory and Language*, 56, 86–102. <http://dx.doi.org/10.1016/j.jml.2006.07.008>
- Wiley, J., Jarosz, A. F., Cushen, P. J., & Colflesh, G. J. (2011). New rule use drives the relation between working memory capacity and Raven's Advanced Progressive Matrices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 256–263. <http://dx.doi.org/10.1037/a0021613>
- Williams, B. A., & Pearlberg, S. L. (2006). Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence*, 34, 177–191. <http://dx.doi.org/10.1016/j.intell.2005.03.007>
- Yntema, D. B., & Trask, F. P. (1963). Keeping track of several things at once. *Human Factors*, 5, 7–17. <http://dx.doi.org/10.1177/001872086300500102>
- Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory mechanisms in attention, memory, and language* (pp. 241–264). New York, NY: Academic Press.

Received September 27, 2018

Revision received January 30, 2019

Accepted February 7, 2019 ■