Journal of Educational Psychology

What's the Best Way to Characterize the Relationship Between Working Memory and Achievement?: An Initial Examination of Competing Theories

Dana Miller-Cotto and James P. Byrnes

Online First Publication, July 25, 2019. http://dx.doi.org/10.1037/edu0000395

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Educational Psychology

0022-0663/19/$12.00

http://dx.doi.org/10.1037/edu0000395

What's the Best Way to Characterize the Relationship Between Working Memory and Achievement?: An Initial Examination of Competing Theories

Dana Miller-Cotto
University of Pittsburgh

James P. Byrnes
Temple University

Although studies have demonstrated a relationship between working memory and achievement in math and reading, it is still unclear *why* working memory might be related to these abilities. In the present article, we examined the viability of several possible theories in 2 separate analyses of math and reading. In each case, we contrasted the predictions of a cognitive filter model, a transactional model, and a positive manifold model using data from the 2011 Early Childhood Longitudinal Study Kindergarten (ECLS-K). Results of path analyses in a structural equation modeling (SEM) framework indicated an excellent fit for the transactional model, while a poor fit was shown for the other 2 models for both math and reading. Findings across these analyses suggested that working memory and achievement interact in a reciprocal, recursive manner over time. Findings are discussed in terms of their implications for theory, practice, and future research.

Educational Impact and Implications Statement

The current study demonstrated support for prior theories that suggested that working memory and prior knowledge work together in an interactive nature over time. That is, the relationship between working memory and achievement may be explained by an individual's ability to retrieve information from long-term memory. This was the case for both math and reading, two domains that have often been seen as very distinct. Findings highlight the importance of understanding the mechanisms underlying the associations between working memory and math or reading ability to improve these skills.

Keywords: mathematics, reading, working memory, path analysis, secondary data analysis

Working memory (WM), or a processing resource of limited capacity involved in the preservation of information while simultaneously processing the same or other information (Unsworth & Engle, 2007), is related to numerous other cognitive abilities employed in everyday life (Diamond, 2013). When creating a mental grocery list, one may need to recall it in the grocery store while also keeping track of the amount of spending money one has as items are added to a shopping cart. In the laboratory, one would use working memory when being asked to listen to a string of numbers and repeat the numbers in reverse, sometimes in the form of an *n*-back task of IQ tests that includes letters or numbers (Conway et al., 2005). In these tasks, the individual is required to first attempt to remember the number or letter string, manipulate the number order, and repeat the numbers aloud in the reverse order (Conway et al., 2005).

In education, it is well-established that WM is correlated both concurrently and longitudinally with mathematics and reading achievement (Byrnes, Miller-Cotto, & Wang, 2018; Peng et al., 2018). Recent meta-analyses have revealed that the average weighted correlation between working memory and math achievement is $r = .35$, and the average weighted correlation between working memory and reading is $r = .29$, respectively (Peng et al., 2018; Peng, Namkung, Barnes, & Sun, 2016). In addition, WM deficits appear to lie at the heart of both mathematical disabilities and reading disabilities (Kudo, Lussier, & Swanson, 2015; Swanson, 2015). Given the correlations suggesting that higher working memory capacity is associated with better math and reading performance, it seems important to understand the nature of this relationship to help those most in need.

Thus, whereas the associations between WM and both math and reading achievement are now beyond dispute, the precise manner in which WM plays a role in these two abilities is still not entirely clear, though several proposals have been advanced (e.g., Perfetti & Stafura, 2014; Raghubar, Barnes, & Hecht, 2010). In the current study, we sought to better understand in what ways WM abilities relate to math and reading as a means of providing insight into the reasons for the similar correlations between WM and both con-

Dana Miller-Cotto, Learning Research and Development Center (LRDC), University of Pittsburgh; James P. Byrnes, Department of Psychological Studies in Education, Temple University.

Correspondence concerning this article should be addressed to Dana Miller-Cotto, Learning Research and Development Center (LRDC), University of Pittsburgh, 3939 O'Hara Street, 811, LRDC, Pittsburgh, PA 15260. E-mail: dcotto@pitt.edu

structs. We also hoped to contribute to ongoing debates about whether WM is domain general or domain specific (Peng & Fuchs, 2016). Theory and research suggest that WM could be related to math and reading achievement in at least three ways. In what follows, we consider the three proposals for math and reading in turn.

The Cognitive Filter Model

Two of these proposals have their roots in classical information processing (IP) theory (Newell & Simon, 1972). In IP theory, the human mind is said to be a limited capacity processing system. The limitations of this system manifest themselves in two ways. One form is reflected in the fact that people cannot process all of the information in a situation, so they selectively encode only some of this material (Newell & Simon, 1972). In more recent formulations, WM is said to act as a kind of cognitive or bottleneck filter that lets in only a subset of the information that is presented to students (Alloway & Alloway, 2010; Robison, Miller, & Unsworth, 2018; Swanson & Fung, 2016; Sweller, 2011). If the amount of information exceeds children's WM capacity because (a) it cannot be "chunked," (b) children fail to limit their attention to just relevant items, or (c) external aids do not reduce the load in some manner, children will fail to process all of the relevant information. In this article, we refer to this proposal as the cognitive filter model.

For instance, proponents of theories aligned with the cognitive filter model, like cognitive load theory (Sweller, 2011), suggest one way to reduce the strain on working memory resources is to learn from worked examples, or a worked-out problem that shows the solution steps. Cognitive load theorists argue that problem solving without such supports taxes one's working memory, increasing the problem's intrinsic load. When the intrinsic load is increased, the learner may fail to attend to the relevant aspects of the problem and will not transfer this knowledge to similar problems that they might encounter (Sweller, 2011). Worked examples are said to reduce the intrinsic load of the problem and allow learning of relevant information. Similarly, one critical factor in determining the effectiveness of problems is their structure. Poorly designed worked examples can also tax working memory resources. For instance, when a worked example uses various sources of information in different formats on the same page (e.g., graphics and text), learning may not occur and can even be deleterious to learning (the so-called split-attention effect; Ward & Sweller, 1990). Atkinson and colleagues (Atkinson, Derry, Renkl, & Wortham, 2000) emphasize integrating features within worked examples to avoid imposing a heavy extraneous load as what occurs with the split-attention effect (Sweller, 2012).

Relationship to Math

Processing limitations are also manifested when people try to solve problems and need to keep directions, the results of substeps, and so on in their minds. Whereas the cognitive filter model pertains to learning or knowledge acquisition, the second proposal suggests that WM could affect performance even after skills are learned as people solve problems that require them to keep directions, goals, and the results of preceding steps in mind. Performance could drop if too much information has to be held in mind (e.g., Hitch, 1978).

Of the two proposals regarding processing limitations, the cognitive load version probably plays a more decisive role across an academic year because students probably do not encounter school tasks in which everything has to be kept in mind (e.g., they can write down steps). If so, cognitive load theory and its variants would predict that students with higher levels of WM can attend to and retain more of the information presented during math lessons that their peers who have lower WM capacity. However, this prediction assumes that teachers do little to reduce cognitive load during lessons. As noted above, there are steps that can be taken to reduce cognitive load during problem solving, such as presenting worked examples and fading, or removing steps in a problem successively, the problem-solving steps (Sweller, 2011; Zamary & Rawson, 2018). When cognitive load is reduced, students can learn how to solve the problems. In any event, in practice, the theoretical distinction between the two kinds of limitations probably does not play out in a typical math lesson because most lessons involve teaching children how to solve computational problems. WM limitations would interfere with the acquisition of declarative, conceptual, and procedural knowledge according to the cognitive filter model and its related corollaries in cognitive load theory (Sweller, Van Merrienboer, & Paas, 1998). For this model, one might envision math knowledge passing through WM, or filtered through WM, to predict later performance, as demonstrated in Figure 1 for the cognitive filter model. As can be seen in Figure 1, there are paths (arrows) extending from both prior WM (e.g., WM1) and prior math (e.g., Math 1) to later math (Math 2), but not from prior math to later WM. However, there is no reason to expect that prior mathematics achievement would affect later WM if a theorist considers it to be a relatively fixed individual difference variable.

Relationship to Reading

A longstanding view of the role of WM in reading is that WM functions as a limited processing space within which the meaning of a sentence can be computed (Caplan & Waters, 1999; Daneman & Carpenter, 1980). When reading times are increased by increasing sentence length, the number of unfamiliar or infrequent words, or grammatical complexity, comprehension suffers, particularly for people with less WM capacity. Stated this way, it can be seen that this account would be the analogue of the cognitive filter model described above for mathematics. Comprehension is demonstrated by asking readers to recall the meaning of a passage that was just read or answer comprehension questions about the passage. If WM capacity keeps people from processing sentences completely, less information would be processed and stored in LTM, or filtered, as suggested above for mathematics. Just as math problems can be constructed to reduce the cognitive load of the

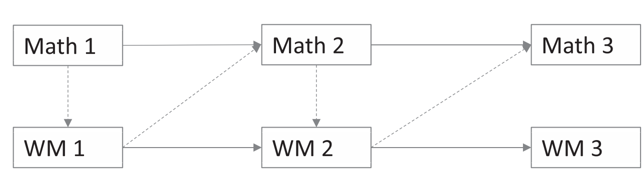

Figure 1. The cognitive filter model. Both WM and prior math determine later math, but prior math does not affect later working memory.

material (e.g., using faded worked examples), texts can be modified to make them more "readable" by making sentences shorter, less grammatically complex, and include only high frequency words (Byrnes & Wasik, 2008). If WM capacity is a relatively stable individual difference variable (as argued in influential works such as Daneman & Carpenter, 1980), then WM would predict concurrent and later reading comprehension. Earlier reading comprehension, however, would not be expected to predict or affect later WM capacity according to this cognitive filter model. Thus, we anticipate that for the cognitive filter performance would be predicted by earlier WM capacity, as demonstrated in Figure 1. But cognitive filter models do not anticipate that prior reading achievement would affect or predict later WM capacity. Once again, WM capacity is viewed as both an individual difference variable and processing resource that affects reading comprehension in the moment. Traditional cognitive filter models do not assume that prior reading performance and practice would affect WM capacity in individual students. As with math, there are paths (arrows) extending from both prior WM (e.g., WM1) and prior reading (e.g., Reading 1) to later reading (Reading 2), but not from prior reading to later WM.

Transactional Model

Besides the cognitive filter model, a second theoretical model that derives from research on the nature of WM proposes that WM tasks not only measure the ability to manipulate, attend to, and store information that is temporarily held in mind, these tasks also provide an index of the ability to retrieve information from long term memory (LTM; Unsworth, 2010; Unsworth, Spillers, & Brewer, 2012). This theory assumes that when the number of items that have to be maintained in WM exceeds some level, the excess items are off-loaded to LTM for later retrieval. This proposal is supported from studies showing a strong correlation between WM and LTM retrieval tasks. As the chief architect of the construct of WM, Alan Baddeley (2010, p. 5) notes:

> . . . it is unsurprising that neuroimaging studies of short-term or working memory tasks also tend to activate areas associated with long-term memory. The crucial question is not whether long-term memory is involved in working memory, but how. In what ways do long-term and working memory interact?

Some findings suggest that students with higher levels of WM capacity develop more robust or distinct representations of material; upon retrieving this material from LTM, they are better able to identify the correct items and avoid confusions (Jones & Macken, 2015). Relatedly, it has been claimed that students with higher levels of WM capacity are less susceptible to interference effects (Hedden & Yoon, 2006; Van Dyke, Johns, & Kukona, 2014). Therefore, the concurrent and longitudinal relations between WM and math achievement tasks may reflect a problem of retrieval from LTM for low WM students rather than (or in addition to) a filtering issue or a problem of the lesser ability to temporarily holding information in mind. Moreover, this LTM proposal is also consistent with the finding that a key diagnostic criterion for children diagnosed with mathematics disability is a developmental delay in the retrieval of math facts (Geary, Hoard, & Bailey, 2012). The linkage between WM and LTM also suggests that retrieval problems for particular items could lessen over time as (a) knowledge slowly accumulates, (b) aspects of math knowledge become better integrated, and (c) the same information is encountered repeatedly (Ericsson & Kintsch, 1995; Jones & Macken, 2015).

Relationship to Math

The transactional model suggests that low WM could affect concurrent performance on achievement tests because children have difficulty recalling the correct information and avoiding interference effects as they are being tested. The opposite would be the case for children with higher levels of WM capacity. However, given at least short-term stability in WM capacity, WM capacity in the fall would also predict math achievement in the spring, but would not reflect only a filtering or bottleneck process. Rather, the predictive relation would reflect consistency in retrieval problems across the academic year. Thus, WM tested at Time 1 would predict math performance at Time 2. Given the cascading effects of knowledge on WM (Ericsson & Kintsch, 1995), however, it would be expected that math achievement at Time 2 might also predict WM performance at Time 3. Thus, there would be both concurrent and cross-lagged relationships over time. In this article, we refer to this kind of model as the transactional model. It shares with the cognitive filter model the idea that processing limitations could affect what is learned in any math lesson. It differs from the cognitive filter model in that the former assumes relatively stable individual differences in WM capacity that solely determine how much children assimilate from classroom experiences over time. According to the cognitive filter model, these individual differences in WM would not be affected by the mathematical knowledge of students (Alloway & Alloway, 2010; Robison et al., 2018; Swanson & Fung, 2016; Sweller, 2011). In contrast, the transactional model assumes that WM performance would be affected by accumulating math knowledge (and vice versa). The transactional model is depicted in Figure 2.

In support of the transactional model, Jones and Macken (2015) investigated whether previous experience affected performance with different types of digit span tasks. Participants were exposed to number spans and digit spans. Sequences were presented more than once to increase participants' familiarity with the sequences. Collective results indicated frequency, not experience, mattered. However, the findings still beg the question of the role of long-term memory on short term memory processes.

As an empirical model then, as demonstrated in Figure 2, we envision WM tested at Time 1 would predict math performance at Time 2, and keeping in mind the cascading effects of knowledge on WM, it would be expected that math achievement at Time 2 might also predict WM performance at Time 3. In other words,

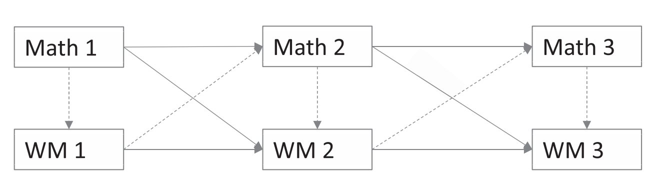

Figure 2. The transactional model. Prior math and prior WM both directly influence each other; but prior math influences later WM and vice versa recursively over time.

prior math and prior WM both directly influence each other; but prior math influences later WM and vice versa recursively over time. This relationship is demonstrated in Figure 2.

Relationship to Reading

An alternative view also exists that represents a synthesis of prior reading and working memory research (e.g., Peng, Wang, Wang, & Lin 2019; Perfetti & Stafura, 2014). In the lexical quality hypothesis (Perfetti, 2007) "word knowledge is paramount to reading skills, suggesting the importance of spelling/vocabulary expertise in reading comprehension (Perfetti & Stafura, 2014). High-quality form knowledge includes phonological specificity, the lack of which has been linked to problems in reading and word learning (Elbro, 1998; Elbro & Jensen, 2005). It also includes orthographic precision." Thus, reading comprehension requires the ability to rapidly retrieve the right pronunciation and contextually appropriate meaning of individual words as they are encountered in a sentence. These meanings should reflect frequency of occurrence (more common meanings are retrieved more rapidly) but frequency should be overridden when the most common meaning is contextually inappropriate (Perfetti & Stafura, 2014). To say that a lexical representation is "high quality" is to say that it allows rapid retrieval of the right meanings and pronunciation, as well as inhibition of incorrect pronunciations and meanings.

Studies have shown that poor readers fail to inhibit the more frequent word meanings and retrieve the less frequent meaning (Perfetti & Stafura, 2014). This finding is highly similar to the point earlier that people with low WM have trouble retrieving the right math facts from LTM and are susceptible to interference effects. Peng et al. (2018) argued that according to the dual processing theory (Evans & Stanovich, 2013) long-term memory determines how WM is employed during a reading task, suggesting that WM's role in retrieval may be more important than reading itself. According to this account, comparing experienced readers and novice readers, novice readers likely need to allocate more of their WM resources to searching in long-term memory and integrating less newly acquired phonological and orthographic representations, representations necessary to perform vocabulary tasks. However, Peng and colleagues (2018) argue that this model does not account for developmental changes between the relationship between WM and reading over time.

If what really lies at the heart of individual differences in WM is the ability to retrieve the right items from LTM and avoid interference effects, a model that combines the lexical quality model and dual processing theory should be seen to be the analogue of the transactional model described above for mathematics. Thus, it would be predicted that WM capacity at Time 1 would predict reading at Time 2, but reading at Time 2 would also predict WM at Time 3. This argument is very similar to Stanovich's (1986) idea of "reciprocal causation" (like the reciprocal nature of Figure 2) in which abilities such as phonological processing and vocabulary predict early reading skills, but reading also tends to improve phonological processing and vocabulary. Demoulin and Kolinsky (2016) make a similar argument and present some suggestive evidence but note that the vast majority of studies have examined the predictive role of WM for later reading, but not the other way around. One study (Ellis, 1990) that tried to examine this predicted cascading relationship found that earlier reading was not a significant predictor of WM using structural equation modeling (SEM), but there were only 40 participants in that study. WM may play a role in the acquisition and immediate retention of information, but once someone is encountering a task, the functioning cognitive variable is likely to be access to relevant knowledge; then, WM has already made its main contribution. When one controls for vocabulary, which is seldom done, WM seems to have less impact. Thus, the transactional model holds promise but has yet to be examined in a direct and methodologically sound way using a large sample.

The Positive Manifold Model

In addition to the aforementioned models, there is also a third possible explanation for the correlations between WM, reading, and math. It is commonly observed in the intelligence literature that many cognitive tasks seem to correlate with each other at a level of around $r = .30$ because each task presumably requires underlying abilities or processes that are common to all of the tests (Burkart, Schubiger, & van Schaik, 2017). The presence of this so-called "positive manifold" often serves as the basis for claims regarding the existence of general or "g" intelligence. Given that recent meta-analyses showed that the average correlation between WM and math performance is approximately $r = .30$ (Peng, Barnes et al., 2018; van der Maas, Kan, & Borsboom, 2014), it is possible that WM neither acts as a filter nor indexes the ability to retrieve items from LTM but, rather, WM and math are correlated because they share some underlying processes or skills (Kovacs & Conway, 2016; Sternberg, 2016). Some of these skills may include allocating attention or keeping substeps to a problem stored and manipulated in WM. Moreover, it should also be noted that many intelligence tests often include measures of WM capacity as subtests. Thus, WM may not promote or facilitate the acquisition or learning of math skills over time. Rather, the correlation that is observed may simply reflect the "g" abilities that are common to WM and math tasks. We call this third possible explanation of the relationship between WM and math, the positive manifold model. It is depicted in Figure 3.

Relationship to Math

The positive manifold assumes that WM and ability share an underlying, common factor. Thus, one might assume that WM and ability would be highly correlated because they might essentially be measuring the same construct. As we demonstrate in Figure 3, we expect that an empirical model would include paths that are highly correlated, with WM and math ability regressed arrows on

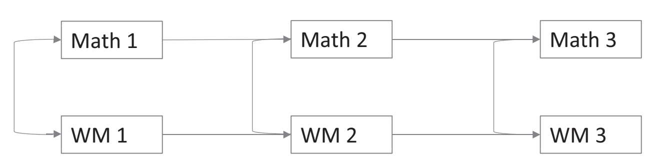

Figure 3. The positive manifold model. Prior math influences later math, and prior working memory influences later working memory, but not cross-lagged effects; working memory and math are correlated concurrently.

to one another because they share an underlying construct, and also predicting both WM and math ability at later time points.

Regarding the literature in this area, Swanson, Jerman, and Zheng (2008) investigated the role of growth in working memory in solving math problems, including skills like assessing problem solving ability, achievement, and students' working memory, inhibition, naming speed, and phonological coding at first grade, second grade, and third grade. Students who exhibited poor working memory showed less growth and lower levels of performance. This was particularly the case if they were deemed at risk for math problem solving difficulties during the first wave. Specific to arithmetic computation, children with poor counting and the inability to remember single digits also tended to have working memory deficits (Siegel & Ryan, 1989).

Relationship to Reading

As was the case for math, there is a third alternative model of the relation between WM and reading. As noted earlier, meta-analyses have shown the average weighted correlation between WM and reading to be $r = .29$ (Peng et al., 2018), about the size of the .30 correlation that emerges in studies demonstrating the positive manifold. The third hypothesis would be, then, that the correlation between WM and reading emerges because these two abilities have some cognitive operations in common such as utilization of the language areas of the superior temporal lobe (Caplan & Waters, 1999). If factor analysis were to be applied to measures of reading and IQ, their loading on the same factor would reflect "g" intelligence. If this third model is true, then one would find only concurrent correlations in a longitudinal study, rather than the mutually predictive, cross-lagged predictors specified in the transactional model.

The Current Study

Recent meta-analyses have shown that the medium-sized association between WM and mathematical and reading abilities is beyond dispute. However, it is still not yet clear why correlations of this size are frequently observed. Peng et al. (2018) and Quinn, Wagner, Petscher, and Lopez (2015) have contributed meta-analyses that include some discussion of the findings according to theoretical accounts. However, the current study differs from these meta-analyses in the following ways. First, a meta-analysis only establishes whether two variables are related and how large the weighted average effect is. When the size of the correlations seem associated with specific moderators, these moderators operate across different studies. In addition, there are no controls for other variables or the ability to test different models of interactions over time. The question of why WM correlates with either math or reading has been vexing cognitive psychologists for years and there have been vigorous debates. The results of prior meta-analyses only partially help to resolve these debates, but more work must be done.

Further, not all of the existing meta-analyses examine the three theories contrasted in this study, neither of the aforementioned meta-analyses theorize why we might expect similar correlations between WM and math and WM and reading. In addition, theories of math ability and reading ability refer to distinct processes (e.g., lexical access and inference making for reading, mental computation and problem solving for math), and cognitive neuroscientific studies show that there is some overlap of function in brain areas but considerable differences as well (Byrnes & Eaton, in press). Thus, we do not think there are a priori reasons for assuming the same variables would predict in exactly the same way. More specifically, we were interested in the following research questions:

1. When comparing the cognitive filter, the transactional model, and the positive manifold model, which model best fits the data?

2a. Is model fit consistent for math and reading?

2b. If the models are indeed different between the domains, in what ways?

Method

Sample

The Early Childhood Longitudinal Study of Kindergarten Cohort, 2011 (ECLS-K: 2011) was sponsored by U.S. federal agencies to provide reliable and comprehensive data on a national sample of children followed from kindergarten to fifth grade using psychometrically sound instruments (Tourangeau et al., 2013). It is a follow-up study to the original ECLS-K study that was conducted between 1998 and 2007. The rationale for the follow-up was that a number of important demographic shifts and government policies had taken place since the 1990s that could potentially have had an effect on several key developmental outcomes (e.g., No Child Left Behind legislation; increase in school choice; increase in immigrant children). Although children from all 50 states were recruited in both studies, neither sample should be considered nationally representative per se because of intentional oversampling of low incidence groups to obtain reliable estimates (e.g., American Indians; Private School students). Across the first three waves of data collection, the number children who had mathematics and reading achievement scores were 17,140 (end of K), 15,103 (end of first grade), and 13,830 (end of second grade). Our analytic sample included 13,480 children who had achievement data in all three grades. Sample sizes at each time point are rounded to the nearest tenths per ECLS guidelines. The demographic breakdown of the sample was as follows: 49% female; 47.6% White, 11.3% African American, 24.6% Hispanic, 8.5% Asian, and 8% "Other" race/ethnic groups. Most (86%) of the children's home language was English.

Measures

Working memory. In the ECLS-K data set, WM is measured using the numbers reversed subtest of the *Woodcock-Johnson III (WJ III) Tests of Cognitive Abilities* (Woodcock, McGrew, & Mather, 2001). This task requires children to repeat back a series of numbers in the reverse order that was said by the tester (e.g., 4, 2 for the sequence 2, 4). Trials started with two digits in a series and then the number of digits is successively increased as the student correctly recalls them in the reverse order in which they were recited. The assumption is that students with more WM capacity are more able to correctly recall these numbers and also manipulate them in the reverse order than their peers with lower WM. ECLS-K data sets used item response theory (IRT) modeling to convert raw scores into WM ability scores in order to account for the fact that more difficult trials are more discriminating with respect to WM (i.e., passing later trials is more indicative of capacity than passing earlier trials). Working memory was assessed in the fall of students' kindergarten school year and again every subsequent year. In the current study, we focused on kindergarten (Year 1), end of first grade (Year 2), and end of second grade (Year 3). Means and standard deviations are presented in Table 1. Regarding the Woodcock Johnson WM measure more generally, scores tend to range from 250 to 550, with an average score being around 450. For interpretation, 10-year-olds tend to score around 500 on this test, whereas a 3-year-olds perform closer to 420 (Schrank, McGrew, & Mather, 2014). For interpretability, quartile scores are presented in Table 2. In the standardization sample and in the ECLS-K sample, the reliability of the WM subtest of the *WJ III* ranged between .90 and .91 for children in Grades K through third.

Math achievement. We included math achievement at the end of kindergarten, end of first grade, and end of second grade. End-of-year mathematics achievement was measured by ECLS-K field agents at kindergarten entry using a standardized mathematics assessment that measured children's conceptual knowledge, procedural knowledge, and problem-solving within specific strands. These strands included number sense, properties and operations; another was patterns, algebra and functions, though the largest proportion of items at all grade levels focused on the "number sense, properties and operations" strand (Tourangeau et al., 2013). Test items were designed to tap a broad range of skills that are typically taught, important skills for that grade, and consistent with national and state standards (e.g., National Assessment of Educational Progress, National Council of Teachers of Mathematics, etc.). Spanish versions of the assessment were administered to children who did not pass the screener test in English. IRT modeling, which was done by ECLS-K data researchers before being made available to the public, was used to create scores that could be compared across time (Tourangeau et al., 2013). A two-stage adaptive test was used to measure these constructs, involving a routing test to ascertain where a student was located developmentally throughout a sequence of items, ending with a second stage of grade appropriate items. This was done to ensure that a student did not take all items while still attaining an accurate score. Regarding the math ability measure more generally, scores tend to range from 6.26 to 95.23 kindergarten, with an average score being around 30. For interpretability, scores by percentile are included in Table 2. The math scale has excellent psychometric properties (α = .93–.94; see Table 1 for means and standard deviations).

Table 2
Correlations Among Math and Working Memory Variables

	Working memory			Math achievement		
	K	First	Second	K	First	Second
WM-K	—	.52	.45	.63	.59	.58
WM-first		—	.54	.54	.58	.57
WM-second			—	.49	.51	.56
Math-K				—	.82	.77
Math-first					—	.85
Math-second						—

Reading achievement. End-of-year reading achievement was measured by ECLS-K field agents at kindergarten entry using a standardized reading assessment that measured children's basic skills (e.g., print familiarity, letter recognition) and sight vocabulary, decoding, vocabulary, and passage comprehension using a variety of literary genres (e.g., poetry, letters, fiction, and nonfiction; Tourangeau et al., 2013). Test items were designed to tap a broad range of skills that are typically taught, important skills for that grade, and consistent with national and state standards (e.g., National Assessment of Educational Progress, National Council of Teachers of Mathematics, etc.). A Spanish version of the assessment was administered to children who did not pass the screener test in English. IRT modeling was used to create scores that could be compared across time and in order to give more weight to more difficult items, which was done by the ECLS-K data team before being made available (Tourangeau et al., 2013). Regarding the reading ability measure more generally, scores tend to range from 21.51 to 90.35, with an average score being around 37.28. For interpretability, scores by quartile are included in Table 1. The reading achievement measure also had excellent psychometric properties after being field tested and constructed by specialists at ETS (alphas = .93 to .95 at each grade).

Analytic Strategy

The analysis was conducted in several stages. Descriptive statistics (i.e., frequencies, skewness, kurtosis) were examined to assess non-normality, outliers, and general data input errors; there was no indication of these found in the data. We then used Mplus to conduct path analyses. Data were missing on less than 1% of

Table 1
Descriptive Statistics for Working Memory, Math, and Reading Achievement

	Working memory			Math achievement			Reading achievement		
	K	First	Second	K	First	Second	K	First	Second
Mean	450	470	481	45.7	67.0	81.2	61.2	83.8	95.9
SD	30.5	25.7	23.2	12.2	15.3	13.8	13.6	15.6	12.3
% missing	.4%	.2%	.1%	.4%	.2%	.1%	.4%	.2%	.1%
25th percentile	426	448	470	36.76	57.20	72.93	52.1	73.4	90.3
50% percentile	448	476	483	45.70	66.64	83.60	59.8	86.3	98.4
75th percentile	476	489	496	53.49	78.30	91.25	68.6	95.8	105

cases, depending on the variable (see Table 1). While it could be argued that such a small amount of missing data would not lead to excessively biased estimates, casewise deletion would lead to a loss of several hundred cases. Further, using casewise deletion is problematic when the data are not missing at random and reducing the *N* can increase the risk of a Type II error. Moreover, the data were not missing at random according to MCAR analyses. Therefore, we elected to correct for missingness using Mplus to maximize sample size and reduce possible bias in estimates (Allison, 2001; Enders, 2010; Rubin, 1987). Although a number of options exists for analyzing structural equation models with missing data, the authors chose full information maximum likelihood (FIML) estimation over multiple imputation (MI) for the following reasons: (a) FIML parameters are unbiased and efficient under MAR, including the more stringent MCAR condition (Enders, 2006); and (b) FIML parameters and standard error estimates are not imputed but estimated directly from the observed data by applying iterative computational algorithms to the sample log-likelihood (Enders, 2006).

To assess model fit, we used the following fit indicators: CFI (comparative fit index), TLI (Tucker-Lewis Index), RMSEA (root mean squared error of approximation), SRMR (standardized root-mean-square residual), and chi square (Kline, 2011). Models that fit the data well are indicated by a CFI $\geq$ .95, RMSEA $\leq$ .06, and SRMR $\leq$ .09 (Hu & Bentler, 1999; Kline, 2011). Good model fit is indicated by a CFI $\geq$ .90, RMSEA $\leq$ .08, and SRMR $\leq$ .08 (Hu & Bentler, 1999). Path models were chosen over structural equation models because there were not multiple indicators of WM, math achievement, and reading achievement in the dataset so it was not possible to derive latent factors. IRT modeling does, however, product ability scores that are comparable to latent factors (Embretson & Reise, 2013).

Results

Analysis I: WM and Math Achievement

We began model testing by first testing the cognitive filter model using path analyses. In the cognitive filter model, later math achievement is influenced directly by prior math achievement (indicated by the solid lines along the top in Figure 1) and indirectly as the function of prior WM (indicated by dotted lines); later WM is only affected by prior WM, as shown in Figure 1. Even after adjusting the model by dropping nonsignificant paths and adding features suggested by modification indices in Mplus, fit for the cognitive filter model was relatively poor concerning the size of the chi-squares, RMSEA, and SRMR with little specific areas of fit improvement, $\chi^2(8) = 3{,}364.51$, $p < .001$, CFI = .938, TLI = .884, RMSEA = .176 (.171–.181), SRMR = .069. The correlations within and between the WM and math outcomes for the 3 years are shown in Table 2.

Next, we examined the fit of the positive manifold model (see Figure 2). This model predicted that while variables within category predict each other at successive time points (e.g., WM at Time 1 predicts WM at Time 2; or math at Time 1 predicts math at Time 2), there are no cross-lagged correlations (e.g., WM at Time 1 predicts math at Time 2). There are, however, concurrent correlations between WM and math, as the positive manifold findings would suggest. Results showed that model fit for the positive manifold was poor even after dropping nonsignificant paths and making adjustments suggested by modification indices, particularly in terms of the size of the chi-squares, RMSEA, and SRMR, $\chi^2(8) = 4{,}886.84$, $p < .001$, CFI = .897, TLI = .821, RMSEA = .213 (.208–.218), SRMR = .184, and there was little indication of where model fit could be further improved using modification indices.

Finally, we considered the fit of the transaction model. To test the transactional model, we modeled direct effects within category as well as cross-lagged effects across time. For both math and WM, later scores are predicted both directly by prior scores in the domain (e.g., Math 2 predicted by Math 1, indicated by paths along the top and bottom) and indirectly as one ability works through the other ability (Math 1 mediated through WM 1's influence on WM 2 mediated by Math 2, indicated by paths). In contrast to the other two models, the transactional model had an excellent fit to the data, $\chi^2(2) = 66.51$, $p < .001$, CFI = .999, TLI = .991, RMSEA = .049 (.039–.059), SRMR = .006, particularly when paths were added between Year 1 math achievement and Year 3 math achievement, as well as between working memory in Year 1 and working memory for Year 3, as suggested by the modification indices noted in Mplus (see Figure 4).

Analysis II: WM and Reading Achievement

The prior analysis of math achievement demonstrated that the best fitting model was the transactional model. To provide further support for this proposal, we considered the generalizability of the results to another domain: reading. The question was whether once again the transactional model would provide a superior fit relative to the cognitive filter and positive manifold models.

The results for the cognitive filter model revealed that the fit was adequate but not the standards for all indices were not met, even after dropping nonsignificant paths and trying theory-consistent modifications suggested by the modification indices: $\chi^2(8) = 3{,}342.70$, $p < .001$, CFI = .989, TLI = .961, RMSEA = .084 (.078–.090), SRMR = .070.

Turning next to the positive manifold model, analyses showed that the fit was even worse: $\chi^2(8) = 4{,}191.03$, p. $< .001$, CFI = .906, TLI = .835, RMSEA = .173 (.169–.177), SRMR = .164.

Finally, we considered the transactional model. As was the case for mathematics, the fit was superior to the other models and the indices met the threshold for an excellent fit: $\chi^2(2) = 82.39$, $p < .001$, CFI = .998, TLI = .988, RMSEA = .048 (.039–.057), SRMR = .012 (see Figure 5).

Discussion

The goal of the present article was to try to gain increased insight into the reasons for the well-established concurrent and

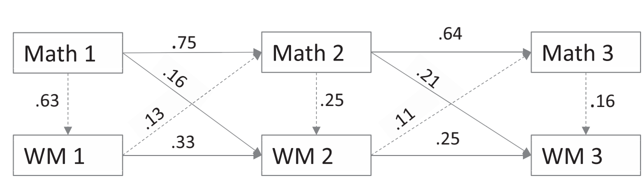

Figure 4. Coefficients for math.

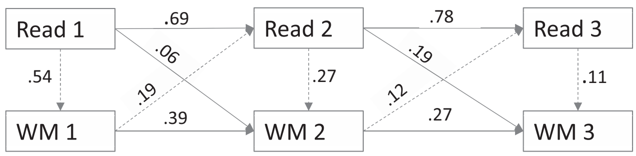

Figure 5. Coefficients for reading.

longitudinal correlations between WM, math, and reading achievement. We considered the viability of three models that have been proposed in the literature: the cognitive filter model, the positive manifold model, and the transactional model. We outline this section based on our research questions below.

Best Fitting Model

The data demonstrated strong support for the transactional model in the analyses for both math and reading. This finding has a number of interesting implications. The first is that the cognitive filter model seems to be partially supported by the data in its characterization of the situational processing constraints that limit how much information students can process when they try to learn math skills or develop fluency in reading. The transactional model has this part of the explanation in common with the cognitive filter model. Students with more WM capacity demonstrated growth in their math and reading skills more than their peers with less WM capacity. However, as shown by the poor model fit, the cognitive filter model fails to acknowledge the mutual, interacting effects of knowledge and WM over time as the transactional model has done herein. This is evident in the finding that whereas both models predict that prior WM capacity would predict how much reading or math skill would be acquired across an academic year, only the transactional model would predict the reciprocal causation construct that prior reading and math skill would predict increases in WM capacity. Students who demonstrated more math skills or reading skills at Time 1 were shown to have more WM ability at Time 2; conversely, students who demonstrated more WM ability at Time 1 were shown to have more reading and math skills at Time 2; and so on.

Second, as Baddeley noted in a quote presented earlier in this article, it is clear that knowledge and WM interact, but it was not clear to him in 2010 *how* knowledge and WM interact (Baddeley, 2010). The present findings provide some interesting clues as to how the interaction may unfold over time. Each time a student gains new knowledge and skill in a domain, this increased knowledge appears related to an increase working memory ability. Increased WM ability, in turn, may help them process more information as they interact with new math problems or develop fluency in reading. Further experimental studies are needed to demonstrate this implied causal relationship. Our findings should be considered strongly suggestive but still correlational.

Implications

Our findings were anticipated by Demoulin and Kolinsky (2016), Stanovich (1986), and Perfetti (1985), but only for reading. Demoulin and Kolinsky (2016) reviewed studies that indirectly supported the idea that reading experience might lead to improved verbal short term memory, but they argued the prediction awaited confirmation in studies that provided direct confirmation. We have not only confirmed this expectation for reading, but also confirmed it for a second domain, math. In one of the few prior studies examining the longitudinal prediction of knowledge to later WM, Ellis (1990) failed to show that prior reading predicted later WM, but as noted earlier, there are reasons to suspect Type II error given that his sample size was rather small ($N = 40$). One of the strengths of our study is that the models were tested using a large, national sample in a longitudinal design. Hence, the findings cannot be dismissed as being a consequence of the idiosyncrasies of a single participating school.

A third implication of our findings pertain to the debates among cognitive psychologists about the nature of WM. The characterization of WM that is most consistent with our results are those of Ericsson and Kintsch (1995), Jones and Macken (2015), and Unsworth et al. (2012). These authors argue that the factor that contributes most to performance on WM is the ability to retrieve information from LTM and overcome interference effects. In our view, this account is extremely similar to the lexical quality account of Perfetti and colleagues (Perfetti & Stafura, 2014). We extend these accounts, however, by showing that WM seems to play a role not only during classroom learning as limiting filtering mechanism, but also when participants are asked to retrieve information on measures of math and reading achievement.

Our findings also provide an alternative explanation of the findings of Peng et al. (2018) who showed in their meta-analysis that the linkage between WM and reading drops out when one controls for both decoding and vocabulary. One way to interpret this finding is to say that the link between WM and reading is spurious. We argue, in contrast, that vocabulary can be properly viewed as a measure of knowledge (or equivalently, an index of lexical quality). Because WM measures are also indicators of retrieval from LTM, controlling for vocabulary would indeed produce a drop in the size of the correlation between WM and reading. In addition, our results may also help explain why training studies of WM seem to show that WM performance does improve after training but there are no transfer affects to achievement (Swanson & Alloway, 2012). If training merely provides or elicits strategies to improve performance on WM tasks, such training would not be expected to lead to achievement effects. Rather, a more effective approach would be to target or increase knowledge, which in turn would lead to improved WM. This prediction awaits confirmation in additional studies.

Finally, our results are consistent with the findings of fact retrieval deficits in children diagnosed with math disability and lexical retrieval deficits in children diagnosed with dyslexia (Geary et al., 2012; Perfetti, 2007). Studies that could provide additional confirmation of this explanation would be ones that carefully examine the responses of children on WM, math, and reading measures that require retrieval to determine the level of intrusion errors and other indicators of susceptibility to interference effects. If the correlation between WM and reading or WM and math drops to nonsignificance when a measure of susceptibility to interference is controlled, such a finding would support the account proposed here.

As can be seen in Tables 2 and 3, the correlations were moderate to strong among all measures. The size of the concurrent correlations between WM and math were comparable in size to that

Table 3
Correlations Among Reading and Working Memory Variables

	Working memory			Reading achievement		
	K	First	Second	K	First	Second
WM-K	—	.52	.45	.54	.56	.54
WM-first		—	.54	.49	.54	.54
WM-second			—	.46	.54	.53
Reading-K				—	.79	.70
Reading-first					—	.86
Reading-second						—

reported in meta-analyses, though somewhat larger perhaps because of the use of IRT modeling for both measures to generate ability scores as opposed to simply number correct. IRT allows a researcher to take in to account the increasing difficulty of the task based on the item characteristic curve; that is, in some cases as the subject's ability increases the probability for a correct response might also increase. These are all taken in to account with IRT and helps us to get to a closer measure of one's true ability. Further, one might expect that the correlations would be larger in the current study, using IRT than in prior meta analyses because simple scores do not account for measurement error. Measurement error as the name suggests reduces the reliability of a measure. Measurement error simply reduces the correlation between two variables; reliability is in essence a correlation with itself, and if the reliability is low it cannot correlate with another variable (Goodwin & Leech, 2006).

Model Fit Between Domains

It is worth noting that findings were similar between math and reading, although in some cases math and reading are seen as distinct abilities. Theories of math ability and reading ability refer to distinct processes (e.g., lexical access and inference making for reading, mental computation, and problem solving for math), and cognitive neuroscience show that there is some overlap of function in brain areas but considerable differences as well (Byrnes & Eaton, in press). Thus, we did not think there are a priori reasons for assuming that the same variables would predict in exactly the same way. However, the findings indicate that this is indeed the case. There are a number of reasons why this might be. The WM measure used herein, although it uses digits, can also be seen as representations much in the same way lexical representations are processed. For this reason, it is unlikely findings would differ.

Limitations

Although the current article adds to the ongoing discussion about WM and mathematics and reading achievement, it is not without its limitations. First, the current study employed the reverse digit span as a measure of WM. It is possible that rather than using this one measure to capture WM, it would have been more useful to use multiple measures of WM to form a composite WM measure. Doing this might give us a better approximation of WM to ensure that we are capturing what we set out to. However, the ECLS-K data set does not provide an alternative measure to WM. It is also worth noting that whereas prior meta-analyses (Peng et al., 2016, 2018) examined WM with respect to specific math and reading skills, our analyses were domain general. Unfortunately, at the time of this writing, ECLS-K did not include these subskills, including a vocabulary measure, and we were unable to include them in our analyses to examine whether such subskills might account for the correlations between WM and math or reading ability. Future work should examine if the transactional model still fits the data as we observe here when considering specific subskills like vocabulary and reading comprehension. Finally, we acknowledge that the data that appear herein are largely correlational. More work is needed to test these theories in an experimental manner.

Future Directions

Some other ways to tease apart the viability of the cognitive filter and transactional models would be through experimental methods (e.g., if children with high and low levels of WM capacity were provided with instructional techniques that help reduce the cognitive load of tasks [e.g., fading of worked examples], such techniques should only be helpful for children with low levels of WM capacity if the cognitive filter model is correct). One might also manipulate material presentation or content difficulty depending on a students' WM ability.

Research in this area may also benefit from examining how material presentation should be altered for students with poor WM capacity. Such a study could be executed longitudinally to examine how these students' trajectories change over time as a result of the presentation format intervention. It may be the case that working memory could be taken into account when designing more effective forms of instruction that improve both WM skills and domain specific skills (Ramani, Jaeggi, Daubert, & Buschkuehl, 2017). Moreover, it is necessary to understand whether the relationship between WM and achievement could have practical implications for interventions targeted at mathematics or reading material presentation (Booth et al., 2017). Subsequent work may be able to design lessons for students that demonstrate low WM capacity, making instructions more specific to individual learner needs. It is well documented that students with different WM deficit profiles may interact with mathematics instruction such that particular instructional strategies would be more effective for students with different WM abilities (Peng et al., 2016). Given the model fit for the transactional model, it may also be the case that instead of focusing on training WM as many studies have attempted to do (for a review, see Schwaighofer, Fischer, & Bühner, 2015) perhaps students should be taught offloading strategies to long term memory when solving math problems or reading. Thus, future work should examine these theories and their predictive power within classrooms or other learning settings.

References

Allison, P. D. (2001). *Missing data*. Thousand Oaks CA: Sage.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106,* 20–29. http://dx.doi.org/10.1016/j.jecp.2009.11.003

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70,* 181–214.

Baddeley, A. (2010). Working memory. *Current Biology, 20,* R136–R140. http://dx.doi.org/10.1016/j.cub.2009.12.014

Booth, J. L., McGinn, K. M., Barbieri, C., Begolli, K. N., Chang, B., Miller-Cotto, D., . . . Davenport, J. L. (2017). Evidence for cognitive science principles that impact learning in mathematics. In D. C. Geary, D. B. Berch, R. J. Ochsendorf, & K. M. Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (Vol. 3, pp. 297–325). Oxford, UK: Elsevier.

Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017). The evolution of general intelligence. *Behavioral and Brain Sciences*. Advance online publication. http://dx.doi.org/10.1017/S0140525X16000959

Byrnes, J. P., & Eaton, J. T. (in press). Educational Neuroscience. In R. Sperling, A. Martin, & K. J. Newton (Eds.), *Handbook of educational psychology and students with special needs*. New York, NY: Routledge.

Byrnes, J. P., Miller-Cotto, D., & Wang, A. H. (2018). Children as mediators of their own development: The case of learning science in kindergarten and first grade. *Journal of Cognition and Development, 19,* 248–277. http://dx.doi.org/10.1080/15248372.2018.1470975

Byrnes, J. P., & Wasik, B. A. (2008). *Language and literacy development: What educators need to know*. New York, NY: Guilford Press.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77–94. http://dx.doi.org/10.1017/S0140525X99001788

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12,* 769–786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466. http://dx.doi.org/10.1016/S0022-5371(80)90312-6

Demoulin, C., & Kolinsky, R. (2016). Does learning to read shape verbal working memory? *Psychonomic Bulletin & Review, 23,* 703–722. http://dx.doi.org/10.3758/s13423-015-0956-7

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64,* 135–168.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. London UK: Psychology Press.

Enders, C. K. (2006). Analyzing structural equation models with missing data. *Structural Equation Modeling: A Second Course, 2,* 493–519.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245. http://dx.doi.org/10.1037/0033-295X.102.2.211

Elbro, C. (1998). When reading is "readn" or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scandinavian Journal of Psychology, 39,* 149–153.

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology, 46,* 375–384.

Ellis, N. (1990). Reading, phonological skills and short-term memory: Interactive tributaries of development. *Journal of Research in Reading, 13,* 107–122.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8,* 223–241.

Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of Learning Disabilities, 45,* 291–307.

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *The Journal of Experimental Education, 74,* 249–266.

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20,* 511–528. http://dx.doi.org/10.1037/0894-4105.20.5.511

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology, 10,* 302–323.

Hu, L. T., & 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

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition, 144,* 1–13. http://dx.doi.org/10.1016/j.cognition.2015.07.009

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Kovacs, K., & Conway, A. R. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry, 27,* 151–177. http://dx.doi.org/10.1080/1047840X.2016.1153946

Kudo, M. F., Lussier, C. M., & Swanson, H. L. (2015). Reading disabilities in children: A selective meta-analysis of the cognitive literature. *Research in Developmental Disabilities, 40,* 51–62. http://dx.doi.org/10.1016/j.ridd.2015.01.002

Newell, A., & Simon, H. A. (1972). *Human problem solving: Vol. 104. No. 9*. Englewood Cliffs, NJ: Prentice Hall.

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., . . . Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144,* 48–76. http://dx.doi.org/10.1037/bul0000124

Peng, P., & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities, 49,* 3–20.

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108,* 455–473. http://dx.doi.org/10.1037/edu0000079

Peng, P., Wang, T., Wang, C., & Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin, 145,* 189.

Perfetti, C. A. (1985). Reading skills. *Psychiatry, 50,* 1125–1129.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18,* 22–37. http://dx.doi.org/10.1080/10888438.2013.827687

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86,* 159–175.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20,* 110–122. http://dx.doi.org/10.1016/j.lindif.2009.10.005

Ramani, G. B., Jaeggi, S. M., Daubert, E. N., & Buschkuehl, M. (2017). Domain-specific and domain-general training to improve kindergarten children's mathematics. *Journal of Numerical Cognition, 3,* 468–495.

Robison, M. K., Miller, A. L., & Unsworth, N. (2018, July). Individual differences in working memory capacity and filtering. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 1038–1053. http://dx.doi.org/10.1037/xhp0000513

Rubin, D. B. (1987). Comment. *Journal of the American Statistical Association, 82,* 543–546.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV tests of cognitive abilities*. Rolling Meadows, IL: Riverside.

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50,* 138–166.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60,* 973–980.

Stanovich, K. E. (1986). Cognitive processes and the reading problems of learning disabled children: Evaluating the assumption of specificity. *Psychological and Educational Perspectives on Learning Disabilities, 10,* 87–131.

Sternberg, R. J. (2016). Groundhog day: Is the field of human intelligence caught in a time warp? A comment on Kovacs and Conway. *Psychological Inquiry, 27,* 236–240. http://dx.doi.org/10.1080/1047840X.2016.1156504

Swanson, H. L. (2015). Intelligence, working memory, and learning disabilities. In T. C. Papadopoulos, R. K. Parrila, & J. R. Kirby (Eds.), *Cognition, intelligence, and achievement: A tribute to J. P. Das* (pp. 175–196). Cambridge, MA: Elsevier Academic Press.

Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook: Vol. 1. Theories, constructs, and critical issues* (pp. 327–366). Washington, DC: American Psychological Association.

Swanson, H. L., & Fung, W. (2016). Working memory components and problem-solving accuracy: Are there multiple pathways? *Journal of Educational Psychology, 108,* 1153–1177. http://dx.doi.org/10.1037/edu0000116

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100,* 343–379. http://dx.doi.org/10.1037/0022-0663.100.2.343

Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation, 55,* 37–76. http://dx.doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J. (2012). Human cognitive architecture: Why some instructional procedures work and others do not. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues* (pp. 295–325). Washington, DC: American Psychological Association.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296. http://dx.doi.org/10.1023/A:1022193728205

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., Hagedorn, M. C., Daly, P., & Najarian, M. (2013). *Early childhood longitudinal study, kindergarten class of 2010–11 (ECLS-K:2011): User's manual for the ECLS-K:2011 kindergarten data file and electronic codebook (NCES 2013–061)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence, 38,* 255–267.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104–132. http://dx.doi.org/10.1037/0033-295X.114.1.104

Unsworth, N., McMillan, B. D., Brewer, G. A., & Spillers, G. J. (2012). Everyday attention failures: An individual differences investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1765–1772. http://dx.doi.org/10.1037/a0028075

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *The Quarterly Journal of Experimental Psychology, 65,* 2397–2410.

van der Maas, H. L., Kan, K. J., & Borsboom, D. (2014). Intelligence is what the intelligence test measures. Seriously. *Journal of Intelligence, 2,* 12–15. http://dx.doi.org/10.3390/jintelligence2010012

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition, 131,* 373–403. http://dx.doi.org/10.1016/j.cognition.2014.01.007

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7,* 1–39.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Itasca, IL: Riverside Publishing.

Zamary, A., & Rawson, K. A. (2018). Are provided examples or faded examples more effective for declarative concept learning? *Educational Psychology Review, 30,* 1–31. http://dx.doi.org/10.1007/s10648-018-9433-y

Received October 3, 2018
Revision received May 31, 2019
Accepted June 12, 2019 ■