The Structure of Working Memory in Young Children and Its Relation to Intelligence

Shelley Gray[a], **Samuel Green**[a], **Mary Alt**[b], **Tiffany P. Hogan**[c], **Trudy Kuo**[a], **Shara Brinkley**[a], and **Nelson Cowan**[d]

[a]Arizona State University, PO Box 870102, Tempe, AZ, US 85287-0102

[b]University of Arizona, PO Box 210071, Tucson, AZ, US 85721

[c]MGH Institute of Health Professions, Charlestown Navy Yard, 36 1st Avenue, Boston, MA, US 02129

[d]University of Missouri–Columbia, 210 McAlester Hall, Columbia, MO, US 65211

Abstract

This study investigated the structure of working memory in young school-age children by testing the fit of three competing theoretical models using a wide variety of tasks. The best fitting models were then used to assess the relationship between working memory and nonverbal measures of fluid reasoning (*Gf*) and visual processing (*Gv*) intelligence. One hundred sixty-eight English-speaking 7–9 year olds with typical development, from three states, participated. Results showed that Cowan's three-factor embedded processes model fit the data slightly better than Baddeley and Hitch's (1974) three-factor model (specified according to Baddeley, 1986) and decisively better than Baddeley's (2000) four-factor model that included an episodic buffer. The focus of attention factor in Cowan's model was a significant predictor of *Gf* and *Gv*. The results suggest that the focus of attention, rather than storage, drives the relationship between working memory, *Gf*, and *Gv* in young school-age children. Our results do not rule out the Baddeley and Hitch model, but they place constraints on both it and Cowan's model. A common attentional component is needed for feature binding, running digit span, and visual short-term memory tasks; phonological storage is separate, as is a component of central executive processing involved in task manipulation. The results contribute to a zeitgeist in which working memory models are coming together on common ground (cf. Cowan, Saults, & Blume, 2014; Hu, Allen, Baddeley, & Hitch, 2016).

Keywords

Corresponding author: Shelley Gray, Department of Speech and Hearing Science, Arizona State University, PO Box 870102, Tempe, AZ 85287-0102 Shelley.Gray@asu.edu.

Working memory is the portion of our human memory system responsible for simultaneously processing and storing incoming information. There are a number of prominent theories of working memory that differ primarily on whether working memory can be divided into domain-specific components, with unique processing and short-term storage capabilities (e.g., Alloway, Gathercole, & Pickering, 2006; Baddeley, 2000; Baddeley & Hitch, 1974; Shah & Miyake, 1996), or whether working memory is part of a larger, more unitary construct primarily guided by the focus of attention (e.g., Cowan, 2001; Engle, 2002). Intelligence encompasses an individual's ability to learn, reason, adapt, understand, and overcome obstacles by thinking (Nesser et al., 1996). Nonverbal intelligence measures assess these abilities using items that do not require overt language, and thus reduce the impact of language ability on performance. In this study we compared the statistical fit of four competing working memory models in children, including a new hybrid model, and then assessed the relationship between our best-fitting working memory models and nonverbal measures of fluid reasoning and visual processing intelligence.

There is an increased interest in the structure of working memory in children because of the central role working memory plays in learning (Alloway, 2009; Alloway, Gathercole, Kirkwood, & Elliot, 2009). In the last decade alone, working memory has been investigated in children with intellectual disability (Van der Molen, 2010; Van der Molen, Henry, & Van Luit, 2014), poor reading comprehension (Caretti, Cornoldi, De Beni, & Romano, 2005), dyslexia (Jeffries & Everatt, 2004), language impairment (Gray, 2006; Leonard et al., 2007; Montgomery & Evans, 2009), autism (Gabig, 2008), attention deficit hyperactivity disorder (Alloway & Cockcroft, 2014), and fetal alcohol syndrome (Paolozza et al, 2014), as well as in children who are learning two or more languages (Blom, Kuntay, Messer, Verhagen, & Leseman, 2014; Morales, Calvo, & Bialystok, 2013). Because working memory is so integral to learning, it is important to determine its structure early in the elementary school years when assessment information can help lead to treatments to prevent future learning problems (Nevo & Breznitz, 2013) and when children are mature enough to complete the wide variety of experimental tasks that permit a full and fair test of working memory structure.

There is also an increased interest in the relationship between working memory and intelligence in children because different components of working memory are thought to predict different aspects of intelligence (Mackintosh & Bennett, 2003) and because some have proposed that working memory actually accounts for individual differences in fluid intelligence, which is the ability to adapt thinking to solve new problems (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Oberauer, Schultze, Wilhelm, & Suß, 2005; but see Gignac & Watkins, 2015).

The Structure of Working Memory in Children

A number of studies have investigated the structure of working memory in children. As shown in Table 1, seven of eight structural studies have considerable overlap in tasks. Although there were differences in the age and primary language of participants, and to some extent how working memory was assessed, results for these modeling studies were quite similar. In general, there was evidence for separate central executive, phonological, and

visuospatial type factors. The exception was the study of 8- to 9-year-old Portuguese children by Campos, Almeida, Ferreira, and Martinez (2013). The fit for their initial confirmatory factor model, with three latent factors (phonological loop, central executive, visuospatial sketchpad), was adequate; however, there was a high correlation (.91) between the central executive and the visuospatial sketchpad factors. They concluded that a model with executive functioning and visuospatial tasks on the same factor was most parsimonious, and therefore they suggested a new two-factor structure as an alternative to the three-factor model. Consistent with this result, Michalczyk, Malstadt, Worgt, Konen, and Hasselhorn (2013) found that a three-factor model fit their data for each age group tested (5–6, 7–9, 10–12), but they reported a "remarkably high correlation between the visual-spatial sketchpad and the central executive" (.81) (p. 227), especially in the younger groups.

Of the studies in Table 1, the investigation by Hornung, Brunner, Reuter, and Martin (2011) is of particular interest because the authors pitted six competing working memory theories against each other in their study of 161 Luxemburgish or Portuguese speaking 5–7 year olds. Using two indicators for verbal simple span, two for verbal complex span, and two for visuo-spatial span, they tested (a) a unitary working memory model, (b) a two-factor model with distinct short-term memory and working memory components, (c) a two-factor model with distinct verbal and visuo-spatial working memory components, (d) a three-factor model (cf. Baddeley & Hitch, 1974) with central executive, phonological loop, and visuo-spatial sketchpad components, (e) a three-factor model (cf. Cowan, 1995a, 1999, 2001) with a domain-general short-term storage component reflecting the focus of attention and two domain-specific components reflecting verbal and visuo-spatial processes, and (f) a three-factor model based on adult research (cf. Unsworth & Engle, 2007) with a common short-term verbal storage component, a working memory residual component representing executive processes, and a general visuo-spatial storage component. The fit for the last three models was excellent and nearly identical, meaning that there was no clear winner. The authors acknowledged limitations in their study, including the need to administer a wider array of tasks. In particular, their battery did not include complex visuospatial tasks or tasks tapping executive function only.

Also missing from the Hornung et al. study, and from most studies of the structure of working memory in children, were tasks designed to assess episodic buffer function. Baddeley (2000) proposed that the episodic buffer is an independent working memory component with its own temporary storage capacity – a kind of 'back-up store that is capable of supporting serial recall, and presumably of integrating phonological, visual, and possibly other types of information' over space and time (p. 419). One study by Alloway, Gathercole, Willis, and Adams (2004) did assess episodic buffer function using two spoken sentence recall tasks. Their final model included episodic buffer, central executive, and phonological loop factors. However, they did not assess visuospatial function; thus, to our knowledge there is no structural test of Baddeley's (2000) four-component working memory model in the research literature.

Open questions about working-memory models

The studies discussed above raise several important questions about models of working memory. First, can the statistical fit of working memory models proposed by Baddeley and Hitch (1974) versus Cowan (1995a, 1999, 2001) be differentiated, provided that a wider variety of indicators are included in the models? As shown in Table 1, we included at least three indicators for each of the four working memory factors studies.

Second, did Hornung et al. (2011) specify their models correctly? According to their representation of Cowan's model, verbal and visuospatial storage were put on equal footing. Cowan actually thought of them differently. In his model (e.g., Cowan, 1988, 1999), the attention-demanding nature of information storage in the focus of attention is postulated for visual information (an assumption now supported by various studies, for example in children by Ang & Lee, 2008; 2010), but attention is largely circumvented when participants can use verbal rehearsal. Cowan has also clearly acknowledged the important role of central executive processes for working memory tasks that require manipulation of information. Therefore, a more accurate three-factor representation of Cowan's model would include as factors (1) the central executive, (2) the focus of attention, and (3) phonological storage and rehearsal.

Third, are there specific findings in addition to statistical fit that would help to adjudicate between the models? There have been reports that visuospatial working memory and central executive function are so closely related that they do not warrant separate working memory factors (e.g., Campos, Almeida, Ferreira, & Martinez, 2013; Michalczyk, Malstadt, Worgt, Konen, & Hasselhorn, 2013). In the Baddeley and Hitch model, neither visuospatial nor verbal storage should be closely related to central executive processes, as they make independent contributions to performance. In contrast with the Cowan model, we would expect a close relationship (though not identical) between the focus of attention, which subsumes visuospatial working memory, and central executive function, given that the executive has control over the focus of attention; but a weaker relationship between these factors and verbal storage in situations conducive to rehearsal, given that rehearsal removes the need for much attention.

Fourth, is there evidence for the existence of an episodic buffer factor as proposed by Baddeley (2000)? There is room for debate about the way to represent the episodic buffer, but one way is to examine situations in which two different kinds of information have to be bound together.

The Relationship between Intelligence and Working Memory in Children

The Cattell-Horn-Carroll (CHC) theory of human intelligence (Carroll, 1993) is a comprehensive psychometric theory of cognitive development, widely accepted as the most empirically supported theory of the structure of cognitive abilities (McGrew, 2005). Because of this empirical support, many intelligence tests are based on CHC theory. Of the 16 broad cognitive abilities described by CHC, seven have been shown to predict academic achievement, and therefore are of primary interest in children: fluid reasoning (*Gf*),

crystallized intelligence, visual processing (Gv), auditory processing, short-term memory, long-term storage and retrieval, and processing speed (McGrew & Wendling, 2010). Of these, Gf has been the focus of working memory researchers because working memory is one of the strongest predictors of Gf in children (Engel de Abreu, Conway, & Gathercole, 2010; Kuhn, in press; Shahabi, Abad, & Colom, 2014; Swanson, 2011; Tillman, Bohlin, Sorensen, & Lundervold, 2009).

In the Hornung et al. (2011) study described above, the authors examined the relationship between components in each of their six tested working memory models by adding a Gf factor to each model. Gf was represented by scores from the Raven's Coloured Progressive Matrices, a nonverbal test of intelligence (Raven, Raven, & Court, 1998). They found that the three-factor model of Baddeley and Hitch (1974) and the three-factor model of Cowan (1995a, 1999, 2001) fit the data best, with nearly identical fit indices. Each of Baddeley's working memory components had correlations of .50 or higher with Gf, but in the Cowan model the component representing shared focus of attention was more strongly correlated with Gf ($r = .58$) than either the domain-specific verbal factor ($r = .24$) or the visuo-spatial factor ($r = .31$). Based on these results the authors concluded that the relation between working memory and Gf was driven by short-term storage because the tasks loading on each factor required storage. This view is consistent with an earlier study by Colom, Abad, Quiroga, Shih, and Flores-Mendoza (2008), who also found that short-term storage was primarily responsible for the relationship between working memory and intelligence in 18–20 year olds, but contrasts with findings in other studies concluding that attention or cognitive control is the primary predictor of fluid reasoning in children (Cowan, Fristoe, Elliott, Brunner, & Saults, 2006; Engel de Abreu, Conway, & Gathercole, 2010).

Purpose of the Present Study

This study had two purposes. The first was to address the unanswered question of whether Cowan's embedded processes model (1995a, 1999, 2001) or Baddeley's three-(Baddeley & Hitch, 1974) or four-component (Baddeley, 2000) working memory models best fit the data for young school-age children. We accomplished this using a wider variety of working memory tasks than previous studies. The second was to assess the relationship between working memory factors, Gf, and Gv. to determine whether short-term storage, attention and cognitive control, or both predict Gf when Gv is also in the model. Given the possibility of the high correlations between central executive and visuospatial factors (Campos et al., 2013; Michalczyk et al., 2013), it was also of special interest to examine the somewhat parallel possibility of a close relationship between Gf and Gv.

The three working memory models we tested are presented in Figure 1. Model 1 represents Cowan's (1988, 1995a, 1999, 2001, 2005) embedded processes model that includes central executive, focus of attention, and phonological storage-and-rehearsal factors. According to Cowan (1988), working memory includes all of the components that are used to hold information temporarily. The core of working memory is the temporarily activated portion of long-term memory that is time-limited and, within it, a focus of attention that can hold several more highly processed, integrated items at once. The central executive processes that are involved in entering information into the focus of attention and initiating mnemonic

strategies also can be considered part of working memory. Early on Cowan (1995b) called the existence of the phonological loop into question, stating that it may "…be just one special application of a more general temporary information storage medium that can contain various types of stimulus features including, at the least, both acoustic and articulatory/phonological features… (p. 5)." The general storage medium to which he referred was the activated portion of long-term memory. Despite emphasizing the potential similarity between different types of activated information, though, Cowan also acknowledged that mnemonic strategies to retain information in working memory may be invoked for verbal information in that covert verbal rehearsal can make memory maintenance somewhat automatic, and thus less reliant on the focus of attention for refreshment compared to other types of information. This would be the case in adults and also in children old enough (i.e., older than about 7 years) to begin to rehearse lists of simple verbal stimuli (Cowan et al., 2005; Flavell, Beach, & Chinsky, 1966; Ornstein & Naus, 1978). Cowan (1995b, p. 7) referred to the study by Guttentag (1984) as evidence that attention is not needed for verbal maintenance in adults or older children nearly as much as in children who have just learned to rehearse. In contrast, nonverbal information that cannot easily be verbally rehearsed appears to require more attention for maintenance (Camos & Barrouillet, 2011). At least in adults, there appears to be adaptive choice between attention and verbal rehearsal as means to retain information in working memory (Camos, Mora, & Oberauer, 2011); this separation between attention and verbal rehearsal should extend to the age group of our study and is quite consistent with the Cowan model.

Model 2 represents Baddeley and Hitch's (1974) three-factor model (with further elaboration by Baddeley, 1986; Baddeley & Logie, 1999) that includes central executive, visuospatial sketchpad, and phonological loop factors. These authors viewed the central executive as an attentional control system (as does Cowan), the phonological loop as a temporary store for speech-based and pure acoustic information that could be refreshed with rehearsal, and the visuospatial sketchpad as a temporary store for visual and spatial information that could also be rehearsed by means of some kind of visual reinstatement (Baddeley, 2007).

Model 3 represents Baddeley's (2000) four-factor model, which added an episodic buffer factor to the previous central executive, visuospatial sketchpad, and phonological loop factors. According to Baddeley (2007) he "proposed to explore the possibility that the executive had a purely attentional role, and was itself incapable of storage" (p. 12), but then needed to account for additional processing capacity observed in tasks that require both memory and processing, especially across different input codes (e.g. visual, auditory). Thus, Baddeley added the episodic buffer "…to form an interface between the three working memory subsystems and long-term memory" (p. 13). The episodic buffer was assumed to have its own temporary storage system and the capacity to bind information from visual, verbal, and perceptual codes with each other and with information held in long-term memory.

With the addition of the episodic buffer, the model of Baddeley (2000) became somewhat similar to that of Cowan (1988, 1999) because Baddeley's episodic buffer took on some of the same qualities as Cowan's focus of attention, including retention of information that is

neither purely phonological nor purely visual or spatial. The models are distinguishable, however, in at least three ways. First, Cowan saw the retention of items that are visual or spatial in nature as dependent on the focus of attention because visual "rehearsal," or refreshment, is assumed not to be semi-automated, even in adults, unlike verbal rehearsal. Therefore, in contrast with Baddeley's models, Cowan's model anticipates a close relation between the central executive components and visual-spatial tasks, the latter being subsumed under the focus of attention. Second, Cowan's model also predicts that it is possible for a verbal stimulus set to be subsumed by the focus of attention when verbal rehearsal is impossible. Such is the case for running digit span, in which digits are presented in a list of unpredictable length; that unpredictability appears to make mnemonic strategies such as rehearsal futile (Cowan et al., 2005; Hockey, 1973) and does not seem to allow much updating, either (Broadway & Engle, 2010; Elosúa & Ruiz, 2008); therefore, this kind of task is quite dependent on attention at the time of recall (Bunting, Cowan, & Colflesh, 2008). Baddeley's (2000) model would not predict that this purely verbal stimulus type would load with visual-spatial tasks under the focus of attention, but rather would predict that its phonological nature would be the overriding characteristic and thus that the task would load on the phonological loop factor. Third, although Baddeley's episodic buffer took on some of the functions of Cowan's focus of attention, they are not the same. According to Baddeley (2007), information that involves the binding of information from diverse sources should load on the episodic buffer. According to Cowan, both types of information should be ascribed to Cowan's focus of attention (e.g., Cowan, 2005). Accordingly, our measures included tasks that could test these differences between the theoretical models.

After determining the best-fitting working memory model we then added subtests from the Nonverbal Scale of the Kaufman Assessment Battery for Children – Second Edition (KABC-2; Kaufman & Kaufman, 2004) to assess the relationship between working memory factors, *Gf* and *Gv*. The addition of *Gv* allowed us to evaluate the differential predictability of the two intelligence factors from the working memory factors.

Our study also allowed a distinction between the roles of working memory storage in general, versus the focus of attention as a storage device. If storage drives the relationship between working memory and *Gf*, we would expect each of the working memory factors to be significant predictors of *Gf* because each includes storage tasks. Conversely, if the focus of attention as a special kind of storage device drives the relationship between working memory and *Gf*, we would expect a stronger relationship between the focus of attention factor and *Gf* than between the central executive and phonological storage and rehearsal factors and *Gf*.

Method

Participants

One hundred sixty-eight 2nd graders (ages 7;0–9;1; years; months) with typical development participated. Children were recruited from public and charter schools that sent consent packets home to all children in second grade. If they wished to participate, parents returned a signed consent form to researchers. Participants in the current study were all children who met inclusionary criteria for typical development (see below) in a larger study of working

memory and word learning that also included children with dyslexia, specific language impairment, dyslexia and specific language impairment, and bilingual Spanish-English speaking children with typical development.

All children in this study met the following inclusionary criteria meant to ensure that children did not have a developmental disability and that their performance was not related to being multilingual: (a) enrolled in or just completed second grade; (b) no history of neuropsychiatric disorders (e.g., ADHD, autism spectrum disorder); (c) no history of special education services; (d) spoke monolingual English; (e) had not repeated a grade; (f) standard score >30th percentile on the Goldman-Fristoe Test of Articulation –2 (GFTA-2; Goldman & Fristoe, 2000) unless scores below that percentile were because of consonant errors on a single sound; (g) standard score >87 on the core language composite of the Clinical Evaluation of Language Fundamentals – Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003); (h) 2nd grade composite standard score >95 on the Test of Word Reading Efficiency, Second Edition (TOWRE-2; Torgesen, Wagner, & Rashotte, 2012); and (i) standard score >74 on the Nonverbal Index of the KABC-II. Descriptive information about children and their performance on inclusionary measures is presented in Table 2. There were 95 girls and 73 boys. Twenty children were Hispanic and 146 Non-Hispanic. Two percent of children were Native American, 2% Asian, 2% Black, 81% White, 12% reported more than one race, and 1% did not report race.

Procedures

Assessment and experimental measures were administered individually in a quiet location, such as the child's school, a local library, the laboratory, or the child's home. The experimental tasks are part of the *Comprehensive Assessment Battery for Children – Working Memory* (CABC-WM; Gray, Alt, Hogan, Green, and Cowan, under development) which includes additional tasks that do not require working memory (e.g., executive function) that are not included in this paper because in this study we were interested in modelling the working memory constructs represented by Baddeley and Hitch (1974), Cowan (1995a, 1999, 2001) and Baddeley (2000).[1] Children also completed word learning tasks not included in this paper. The working memory tasks were presented in a computer-based, pirate-themed game that took six to seven 2-hour sessions to complete over a period of approximately two weeks.

Children were seated 52 cm from a touchscreen computer monitor next to a research assistant (RA). A green circle, which was the resting position for the child's response hand,

---

[1]Non-memory tasks designed to examine executive function were omitted from the model because it is a model of working memory, but theoretically one might expect these tasks to correlate with the memory-based central executive tasks, given that they presumably rely upon similar mental processes to control memory versus other cognitive functions. To begin to assess these relations, we first determined each participant's factor score for the memory-based central executive tasks using principal components analysis. This score was found to correlate with accuracy in a stop-signal task, $r(131) = .30$, $p<.001$. No other correlation with an originally-intended dependent measure was larger than .18 among two types of task-switching procedures and two types of Stroop procedures (unimodal and cross-modal). One might attribute the absence of robust correlations for these four tasks to the fact that the intended dependent measure in each case involved a subtraction of one condition mean from another, resulting in some low task reliabilities (for details see Cabbage et al., 2016, submitted). Correlations were higher for non-subtracted means [notably, Global-Local classification same and switch accuracy, $r(115) = .26$ and .27, p's <.01; sorting monsters by pattern or color, $r(130) = .28$ for same-task and .37 for within-block task-switch accuracy, p's 001; Stroop conflict trial response times, $r(127) = .33$, $p<.001$]. These correlations may be theoretically valid given that the issue of whether there is a conflict or not is relevant in both non-switch and switch trials and in both Stroop and control trials (cf. Kane & Engle, 2003).

was secured on the table directly in front of the computer at a distance of four inches from the monitor. Children were instructed to use their dominant hand to respond to all tasks requiring a touch screen response or key-press. Between trials children were instructed to return their hand to the green circle. Children and RAs wore headsets with integrated microphones used to time and record children's verbal responses.

Experimental Tasks

We administered 13 different experimental working memory tasks. The order of administration was randomized across and within research sessions. A general description of each task is included below with a more detailed description in the Appendix. Children began the series of games by selecting their own pirate avatar who traveled from island to island (day to day) with the child and earned rewards for participating in the games. Each task began with instructions delivered by a guide pirate on the computer screen, along with a demonstration of how to play the game. The guide pirate enlisted the help of the child to solve a problem, thus providing motivation for task completion. This was followed by training trials that varied by task (see Appendix). Children were required to pass the training trials to proceed to the game. If they did not pass training the pirate guided them to the next game. Children did not know whether each of their individual responses was correct or incorrect; however, at the end of the game they received a virtual pile of rocks and gold coins that reflected their overall performance. This was included for motivational reasons. Children enjoyed spending their gold coins on their own pirate avatar at a virtual pirate store at the end of each research session.

Central Executive Tasks—Central executive tasks were designed to assess working memory using visual and auditory tasks that required storage and manipulation. To successfully complete the tasks children had to maintain activated memory representations while processing incoming information.

N-back Auditory: This task was presented in the context of a robot band playing different instruments, which were pure tones that varied in frequency (1000Hz, 1250 Hz, 1500 Hz, 1750 Hz, and 2000 Hz). Children saw the still image of a robot band and listened to a series of tones. Their task was to decide whether a new tone was the same or different from the previous tone in the sequence. After each tone presentation the robot band image disappeared and was replaced by a green rectangle response cue that remained on the screen for 3000 ms. Children responded by pressing a designated key on the keyboard labeled with a green sticker for 'same' or a red sticker for 'different.' The next trial began immediately after the child's response or after the 3000 ms response period ended. Accuracy was recorded by the computer.

N-back Visual: This task was presented in the context of robots playing a game with patterned game pieces. Each game piece was a black square with different patterns of white dots. Children were shown a series of individual game pieces and, after each piece was shown, asked to judge whether the pattern was the same or different from the preceding piece. Each trial began with the presentation of the game piece that remained in the center of the screen for 1000 ms. after which the piece disappeared and a blank response cue screen

appeared. Children responded by pressing a designated key on the keyboard labeled with a green sticker for 'same' or a red sticker for 'different.' After the child's response or 3000 ms. had elapsed the next trial began. Accuracy was recorded by the computer.

Number Updating: Inspired by adult work relating updating to working memory and the focus of attention (e.g., Oberauer, 2002), this task was presented in the context of making yoyos and teddy bears at a toy factory. Children were asked to remember how many of each type of toy needed to be made to fill a toy order. Each trial began with two squares rimmed in black displayed on the screen, one with yoyos in the background and the other with teddy bears. Each square contained a single digit that remained on the screen for 2000 ms. The squares were then replaced by operation squares rimmed in red showing an addition operation (e.g. + 1) on one of the squares to indicated that the child should add that many yoyos or teddy bears to the running total for that toy. The operation squares remained on the screen for 500 ms, after which blank squares (with yoyos and teddy bears in the background) rimmed in green appeared to cue the child to provide a verbal response reporting the updated running totals for each type of toy. The child reported the two numbers which were entered by the RA using the keyboard. After the child's response was entered the next trial began after a 50 ms interval. To score a 1 for the trial the child had to report correct running totals for both toys. If the child responded with an incorrect number, but used that number from that trial forward to correctly report the running total, they scored a 0 for the initial incorrect trial, but received credit for subsequent correct trials.

Short-Term Phonological Memory Tasks—These tasks were designed to assess phonological short-term storage capacity with minimal reliance on lexical or semantic knowledge. In the running task the child did not know how many items would be presented. This unpredictability is thought to reduce the ability to group and rehearse items (Cowan et al., 2005).

Digit Span: The goal of the game was to play copycat with robots who read lists of numbers with digits 1–9 (excluding 7 because it has two syllables) in random order. Each trial began with the auditory presentation of a series of numbers, after which a green rectangle appeared on the screen to prompt the child to verbally recall as many numbers as possible in sequence. The RA wrote down the child's responses then entered them into the computer using the keyboard. Verbal responses were also audio recorded by the computer.

Digit Span Running: The goal was to play copycat with sea monsters who spoke lists of numbers with spans from 7–10 digits in length. The procedures were the same as Digit Span, except that children did not know how many digits would be presented and they were asked to recall as many numbers as they could from the end of the list in forward order.

Nonword Repetition: This task was presented in the context of a pirate building a bridge over a river. As children repeated each nonword an additional piece of the bridge was added until the bridge was complete. Each trial began with the auditory presentation of a nonword, after which the child verbally repeated what was heard. After the child made a verbal attempt the RA advanced the program to the next trial. Children's responses were audio recorded by the computer for later scoring in the lab by trained phonetic transcribers.

The 16 nonwords (four each at 2-, 3-, 4-, and 5-syllable lengths) each contained low frequency biphones and had no phonological neighbors. Nonwords of the same syllable length did not differ statistically in spoken duration. Children scored 1 point for each nonword with all consonants repeated correctly. If a child demonstrated consistent articulatory substitutions (e.g., /s/ for /z/) as evidenced by their performance on the GFTA-2, they were not scored as incorrect. Inter-rater transcription reliability for phoneme-by-phoneme consonant scoring was 87%.

Short-Term Visuospatial Memory Tasks—These tasks were designed to assess children's short-term memory for visual information, including shapes and locations that could not easily be remembered using verbal labels. In the running tasks the child did not know how many items would be presented.

Location Span: The goal of the game was to remember where a series of arrows pointed to help direct the pirate to buried treasure. Each trial began with the appearance of a black dot in the center of a white screen. At each span length a sequence of black arrows appeared for 1000 ms each, one at a time, pointing to a discrete location radiating out from the center at 8 equidistant angles. After the entire sequence of arrows had been shown, eight red dots appeared in a circular pattern around the screen (but not at locations typical of a clock face) to show all of the possible locations where arrows could point. Children were asked to touch a red dot for each location on the screen where an arrow had pointed and to do so in correct order. The dots remained on the screen until the child had touched the correct number of dots, after which the next trial began. Children were allowed as much time as necessary to make their selections.

Location Span Running: This task was the same as the Location Span Task except that children did not know how many locations would occur in the series (which was randomized by the computer). They were asked to touch the red dots indicating their recall of as many locations as they could remember from the end of the list in forward order. When finished children touched a 'NEXT' button on the screen to begin the next trial.

Visual Span: The goal of the game was to help the pirate remember which 'gems' (black polygon shapes) appeared on the screen in the correct order. A single polygon appeared in the center of the computer screen, remained on the screen for 1000 ms, and was replaced by the next polygon. At the end of the trial a selection screen appeared with empty response boxes equivalent to the number of polygons in the sequence. From a field of six available polygons, children were asked to touch the polygons they had seen in the order in which they had appeared. When children had selected the final polygon in the sequence, the next trial began. Children were allowed as much time as necessary to make their selections. A trial was scored with a 1 if the entire span length was correct or a 0 if one or more items were incorrect.

Visual Span Running: The Visual Span Running task was similar to the Visual Span task except that children did not know how many polygons would occur in the series (this was randomized by the computer). Children were asked to recall the polygons in order when prompted. At the end of the sequence all six polygons were displayed on the screen and

children used the touchscreen to select as many polygons as they could remember in the correct order.

Binding Tasks—These tasks were designed to assess working memory capacity when two different types of stimuli were presented within the phonological or visuospatial domains, or across the phonological and visuospatial domains, that had to be held together in working memory to respond correctly to the task.

<u>**Phonological Binding Span:**</u> The goal of the game was for children to remember nonword to sound pairings in the context of robots ordering candy at a candy store using non-speech sounds (e.g., mechanical noises, beeps) that named the candy. The pairings differed for each span presented. Nonwords had low phonotactic probability (7–13 neighbors each) and were drawn by the computer from a pool of 11 single-syllable CVC words. No sound or nonword was repeated within a trial. Each trial began with the appearance of a robot on the screen that stayed on the screen while the nonspeech sound was presented. After 500 ms the nonword was presented while the robot image remained on the screen. After 2000 ms a speaker icon appeared on the center of a white screen while the nonspeech sound was presented. A green rectangle appeared on the screen to prompt the child to say the nonword that had been paired with the nonspeech sound that was just played. After the child responded a research assistant advanced the program to the next trial. The number of nonspeech sounds and nonword pairings in each trial varied from one to four. Responses were recorded by the computer for later scoring in the lab by trained phonetic transcribers. A nonword was considered correct if all consonants were produced correctly. Consistent articulatory substitutions were not counted as incorrect. Inter-rater transcription reliability was 94%.

<u>**Visual-Spatial Binding Span:**</u> The game presented a 4×4 grid with 16 squares on the screen. For each span a polygon was displayed in a discrete location on the grid for 1000 ms followed by a blank grid for 500 ms then a new polygon was displayed in a different location on the grid for 1000 ms. Up to six polygons were displayed in a sequence depending on the span length for that particular trial. After the last polygon in the trial was displayed a blank grid appeared on the screen next to a field of the six polygons. Children were instructed to use the touchscreen to select and drag polygons to their proper locations within the grid in the sequence they appeared. A trial was scored with a 1 if the entire span length was correct or a 0 if one or more items were incorrect.

<u>**Cross-Modal Binding:**</u> The game was to remember the nonword names for black polygon game pieces. For each series a polygon appeared by itself on the screen with a simultaneous presentation of its name through headphones. Nonwords were dissimilar from each other, which meant that they did not contain the same vowels. They had low phonotactic probability and neighborhood density. Each trial varied in the number of stimuli presented, ranging in span length from one to six polygons. After the presentation of the last stimulus in a span, a selection screen showing the field of all six polygons appeared. Children heard each nonword and used the touchscreen to indicate the polygon that had been paired with that particular nonword. The nonwords were not replayed in the order in which they had

been presented. A trial was scored with a 1 if the entire span length was correct and with a 0 if one or more items were incorrect.

Task Reliabilities

We calculated the reliability of our working memory tasks by calculating split-half and split-third coefficients, which are special cases of the more general K-split coefficient (Green & Yang, 2005; Raju, 1977). Reliability for each task is shown in Table 3. A detailed description of the use of internal consistency coefficients for estimating reliability of experimental task scores may be found in Green et al. (in press). Most reliabilities were moderate to high in value.

Intelligence Measures

We administered four nonverbal subtests of the KABC-II. According to the test manual, internal consistency estimates for factor scores range from .88 to .93 (Kaufman & Kaufman, 2004). The technical manual classifies Block Counting and Triangles subtests as *Gv* measures and Pattern Reasoning and Story Completion subtests as *Gf* measures. These classifications have been confirmed by an independent study of children ages 6–16 years (Reynolds, Keith, Flanagan, & Alfonso, 2013).

Statistical Modeling Methods

Confirmatory factory analysis (CFA) was employed to assess the fit of Cowan's three-component embedded processes model, Baddeley and Hitch's (1974) three-component model, and Baddeley's (2000) four-component model. In Figure 1, we show the relationship between the hypothesized factors and the 13 task variables for each of these models.

In addition to the CFAs, we used structural equation modeling to clarify the relationships between the 13 working memory tasks and the four KABC-II nonverbal intelligence subtests. The correlations among these tasks and subtests are presented in Table 4. For these analyses we specified the structures for the working memory tasks that were consistent with those for the CFA models, as well as *Gf* and *Gv*, the two intelligence factors underlying the four KABC-II subtest scales.

All analyses were conducted with Mplus 7.2 (Muthén & Muthén, 1998–2012) using maximum likelihood parameter estimation with standard errors and chi-square test statistics that are relatively robust to non-normality (MLR). In addition, MLR allowed for missing data. Model fit was assessed globally using three statistics: the χ^2 test statistic, the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). Rejection of the null hypothesis based on the χ^2 implies a lack of support for the hypothesized model. The CFI compares the fit of the hypothesized model to a null model and ranges in value from 0.00 to 1.00, with higher values indicating better fit. A traditional cutoff value for good fit with the CFI is .90 (Bentler, 1990). The RMSEA is an absolute index that yields the value of 0.00 if the model fits the data perfectly and increases in value with poorer fitting models. Browne and Cudeck (1993) suggested the following cutoffs for RMSEA: .10 or less for adequate fit, .08 or less for reasonable fit, and .05 or less for close fit.

The hypothesized models were not nested within each other, and consequently chi square difference tests were not conducted to assess relative fit of models. We assessed relative fit by comparing their CFIs and RMSEA, as well as the Akaike information criterion (AIC). The AIC is useful in selecting between models while taking into account model complexity.

Besides assessing models based on global fit, we also examined standardized and raw parameter estimates to ensure that their values were reasonable, standardized residual covariances to evaluate lack of fit for particular sample covariances, and Wald and Lagrange multiplier tests to determine whether particular parameters should be fixed to zero or freely estimated.

Results

Working Memory Models

We initially evaluated the three working memory models shown in Figure 1. The Cowan embedded process model and the Baddeley and Hitch (1974) models converged to solution without estimation problems. The Baddeley four-component model converged to solution, but the correlation between the visuospatial sketch pad factor and the episodic buffer factor was 1.08, an out-of-bounds estimate. These results indicate that an episodic buffer factor is empirically indistinguishable from the visuospatial sketch pad factor. The episodic buffer factor is the primary feature of the Baddeley four-component model, and thus the data failed to support this model.

As shown in the top part of Table 5, both the Cowan embedded process model and the Baddeley and Hitch (1974) models evidenced good fit. However, the Cowan model showed better fit on all fit indices. For example, the AIC was 6 points lower for the Cowan embedded process model than for the Baddeley and Hitch model.

As shown in Figure 2, the specification of the two models differed with respect to loadings on three indicators: cross-modal binding, digit span running, and phonological binding. To develop a better understanding of the fit of these models, we specified a combined (hybrid) model that was identical to the models by Cowan and by Baddeley and Hitch (1974), except we allowed for cross-loadings between the last two factors and three indicators: cross-modal binding, digit span running, and phonological binding. The combined model - the third model in Figure 2 - fit quite well and similarly to the Cowan model. For example, the AIC for both the Cowan and combined models was 7436.

Given the overall fit, we next examined the local fit of the combined model. The results indicated that the phonological binding task was a function of the "phonological" factor, which supports both the Cowan and Baddeley and Hitch (1974) models, but it was nonsignificantly and minimally related to the "focus of attention/visuospatial sketch pad" factor, as specified in the Cowan model. Also, the results suggested that the cross-modal binding task was a function of the second factor (focus of attention in Cowan's model and visuospatial sketch pad in the Baddeley-Hitch model), which supports both models, but it was nonsignificantly and minimally related to the "phonological" factor, as specified in the Baddeley and Hitch model. Finally the digit span running task was significantly and

moderately related to the second factor, consistent with the Cowan model, but nonsignificantly and less strongly related to the "phonological" factor, which fails to support the Baddeley and Hitch model. Overall, the results show greater support for Cowan's embedded processes model; however, not surprisingly, the fit of the two models are not dramatically different because they have similar structures.

Working Memory and Intelligence

To assess the relationship between working memory, *Gf*, and *Gv*, we incorporated the four KABC-II intelligence subtests into each of the best-fitting working memory models. We began by specifying one of the two working memory models and then included the two intelligence factors underlying the four KABC-II subtests. The intelligence factors were a function of the three working memory factors for each working memory model, and the covariance between the disturbance terms for the two intelligence factors were freely estimated to minimize misspecification. These models are presented in Figure 3.

Both models converged to solution and yielded in-bound estimates. As shown in the bottom portion of Table 4, the models fit adequately, although not as well as the models for the working memory tasks alone. Examining the residuals matrices and the modification indices, it was apparent that the lack of fit was primarily due to the relationships between the working memory tasks and the nonverbal intelligence tasks. More specifically, additional fit could be gained by incorporating covariances between the residuals of the working memory tasks and the intelligence subtests. However, the inclusion of these covariances would be post hoc and not theory driven. Because the fit of the models was adequate, we proceeded with interpretation of the results.

When all three working memory factors were allowed to predict the intelligence factors, only the focus of attention factor was significantly related to the two intelligence factors for the Cowan model (in keeping with Cowan et al., 2005). Similar results were found with the Baddeley and Hitch (1974) model; the visuospatial factor was most strongly related to the two intelligence factors, although the only significant path was between the visuospatial and *Gv* factors.

The reported standardized coefficients for these models indicated the effect of each of the working memory factors on the intelligence factors, partialling out the other working memory factors. Thus, we also computed the correlations between the working memory factors and the intelligence factors, as presented in Table 5. For the Cowan model the correlations of the focus of attention factor with the *Gf* and *Gv* factors were significant. In addition, the correlations of the central executive factor with the intelligence factors were significant. As expected, the correlations between the phonological storage and rehearsal factor were nonsignificantly and very weakly related to the intelligence factors.

Similar results were found for the Baddeley and Hitch (1974) model. The central executive and visuospatial sketch pad factors were significantly related to the intelligence factors. The correlations with the phonological loop factor were somewhat more strongly related to the intelligence factors for the Baddeley and Hitch model in comparison with the Cowan model; however, none of these correlations were significant.

Discussion

The first purpose of this study was to determine whether Cowan's embedded processes model (1995a, 1999, 2001) or Baddeley's three- (Baddeley & Hitch, 1974) or four-component (Baddeley, 2000) working memory models best fit the data for young school-age children, and what constraints might be placed on the models that worked well. The second purpose was to assess the relationship between the best fitting working memory models, *Gf,* and *Gv*.

The Structure of Working Memory in Young Children

We first asked whether the statistical fit of working memory models proposed by Baddeley and Hitch (1974) versus Cowan (1995a, 1999, 2001) could be differentiated, provided that a wider variety of indicators were included in our models than in previous studies including Hornung et al. (2011). We found that Cowan's three-factor embedded processes model fit the data better than Baddeley's three-component model. The structures of both models were very similar, but with important conceptual differences in what the second factor represents - the focus of attention or the visuospatial sketchpad. This differentiation relies primarily on loadings for the digit span running task.

Cowan considers performance on running memory tasks to reflect the capacity of the focus of attention because the unpredictable endpoint of the list interferes with rehearsal (e.g., Crowder, 1969; Hockey, 1973); therefore, in Cowan's view the number of items recalled reflects the number that can be maintained within the focus of attention (Bunting, Cowan, & Saults, 2006; Bunting et al., 2008). For this reason, Cowan's model placed the digit span running indicator on the focus of attention factor (cf. Cowan et al., 2005). In contrast, it is reasonable to assume that Baddeley and Hitch (1974) would place digit span running on the phonological loop factor because it is entirely verbal in nature and Baddeley and Hitch did not credit the central executive with its own storage capacity.

Our analyses showed that the digit span running loading was significant on Cowan's focus of attention factor and on Baddeley and Hitch's phonological loop factor; however, in the comparison model (Figure 2) when digit span running was allowed to load on both factors, it was significant only on the focus of attention factor, indicating that attention appeared to be more important for task performance than the phonological nature of the task.

It is important to note that digit span running (which involves spoken stimuli) has historically been of special theoretical importance for the Cowan model (see, for example, its role in Cowan et al., 2005, and cf. Bunting et al., 2008). Unlike most measures with phonological stimuli, rehearsal and grouping are impeded in running span despite the use of speech materials because of the unpredictability of the list termination point. Therefore, it is indeed of special, previously-anticipated importance that this measure loads not with the other phonological measures, but with visual measures, which all seem to require more attention than the other language-based, simple span measures we used.

The Episodic Buffer—We evaluated whether there was evidence for the existence of an episodic buffer factor as proposed by Baddeley (2000). We tested his four-factor model with

central executive, phonological loop, visuospatial sketchpad, and episodic buffer factors. Indicators for the episodic buffer included one task to assess within-domain binding for speech (binding speech sounds to nonspeech sounds), one to assess within-domain visuospatial binding (binding shapes to spatial locations), and one to assess cross-domain binding (binding shapes to nonwords). When we fit the four-factor model, we found that the central executive and visuospatial sketchpad factors could not be empirically differentiated, indicating that a four-factor model was not viable. That is not to say that the functions specified for the episodic buffer by Baddeley are unimportant, but that its proposed binding functions do not appear separate from the other functions of the focus of attention as a storage device.

Our results differ from the only modeling study to date that assessed episodic buffer function in children. Alloway, Gathercole, Willis, and Adams (2004) used sentence recall tasks to assess the episodic buffer, reasoning that to repeat sentences children must integrate information from temporary memory subsystems to remember the exact words in sequence along with the products of semantic and syntactic analysis. They found evidence for an episodic buffer factor using these tasks, but noted that the episodic buffer factor was highly correlated with both their phonological loop (.77) and central executive (.66) factors (but note only .36 with the visuospatial sketchpad factor). We argue that our binding tasks provided a more stringent test of the episodic buffer because in addition to verbal binding tasks, we also assessed visuospatial and cross-modal binding. The latter is particularly important because in his (2007) description of the episodic buffer, Baddeley emphasized the need for a mechanism that allows 'verbal and visuospatial subsystems to interact with each other, and with long-term memory' (p. 147).

Based on our results, it is not clear that binding requires a special kind of storage such as the episodic buffer. As shown in Table 4, our visuospatial binding and cross-modal binding tasks both showed their largest correlations (all above .30) with various visuospatial tasks, whereas the phonological binding task showed its largest correlation with nonword repetition (.27) and no other correlations above .21. From the perspective of Cowan's model, the first two kinds of binding worked well being subsumed under the focus of attention, whereas the phonological binding task may have taken advantage of phonological rehearsal and loaded only with that factor, not with the focus of attention as anticipated.

Relationship among Working Memory Factors—Factors identified as statistically separate (with a loss of fit if any factor is excluded) are still expected to be related to one another for various reasons, much as heart and lung functions are related. In the Cowan working memory model, a stronger relationship would be expected between the central executive and focus of attention factors because of their mutual reliance on attentional processes. The relation between these two factors and the phonological storage and rehearsal factor would be expected to be lower because the ability to rehearse decreases the need for attention. In contrast, in the Baddeley and Hitch working memory model, a similar relationship would be expected between the central executive and the phonological loop and visuospatial sketchpad factors because these two slave systems are under the control of the central executive, but a weaker relationship between the two slave systems is expected because they operate relatively independently.

The pattern of relationships among factors favors the Cowan model, with a strong relationship between the central executive and focus of attention factors and weaker relations between the central executive and phonological storage and rehearsal factors, and between the focus of attention and phonological storage and rehearsal factors. The same pattern of relations was found for the Baddeley and Hitch model, with the relationship between the visuospatial sketchpad and the phonological loop being even stronger than in the Cowan model, likely due to the digit span running task loading on the phonological loop factor in the Baddeley and Hitch model.

These results speak to another question regarding whether visuospatial working memory and central executive functions are so closely related that they may not warrant separate working memory factors, as suggested by Campos, Almeida, Ferreira, and Martinez (2013) and Michalczyk, Malstadt, Worgt, Konen, and Hasselhorn (2013). In these studies the correlations between the central executive and visuospatial factors were .91 and .81 respectively, but in our study the correlation was lower in both the Cowan (.75) and Baddeley and Hitch (.76) models. We provided a more comprehensive test of these relationships because our central executive tasks included both visual and verbal tasks. In addition, we also included both verbal and visual running tasks, which are known to tax attention. We conclude that although the central executive and focus of attention are highly correlated, they are distinct and warrant separate factors.

Convergence of Baddeley and Cowan Working Memory Models—As indicated by Hornung et al. (2011), our own findings, and other research studies the Baddeley and Cowan working memory models appear to be coming closer together. Perhaps the convergence started with the introduction of the episodic buffer in 2000, which was intended to handle some of the same kinds of phenomena that Cowan relegated to the focus of attention. At one point, though, Baddeley and colleagues proposed that the episodic buffer did not require attention, but rather retained bindings automatically (Allen et al., 2006). We can rule out this version of the episodic buffer because our SEM models showed that there was no separate faculty for binding tasks, which generally loaded with the visual tasks or the focus of attention. Recently though, Baddeley and colleagues have found evidence for a focus of attention (Hu et al., 2016), so they are coming closer to the Cowan (1988, 1999, 2005) model. Cowan et al. (2014) also allowed that more information is off-loaded out of the focus of attention than was previously thought, so the models are beginning to converge in postulating both attention-demanding and attention-free types of information storage. Like Cowan (1988), Hu et al. described a process whereby the direction of the focus of attention was determined partly by salient features of the environment that are beyond the participant's control (in their case, a list-final visual item that did not need to be recalled along with the rest of the list of colored shapes) and partly by voluntary, central executive processes (in their case, responsive to a payoff system that favored some serial positions above others).

Given this convergence, we propose that our combined model (Figure 2) that included links from the Baddeley and Cowan models could serve as a test case for future working memory modelling experiments. The combined model captured the variance as well as the Cowan model. We realize that researchers who prefer the Baddeley or Cowan models for theoretical

reasons may take issue with the combined model. In this case our results still provide the best comparison of these models to date and they help to limit the version of each model that can be supported. Specifically, the Baddeley model cannot be the one with an autonomous, non-attention-using episodic buffer to retain bindings; the retention of bindings must be folded in with other attention-demanding retention. The Cowan model cannot be a version that puts the focus of attention on a pedestal by itself; it must be one in which the pool of resources for manipulating items (the central executive) is to some degree separate from the pool of resources for retaining items (the focus of attention) and one that acknowledges a special status for phonological materials of a list length known to the participant.

The Relationship between Working Memory, Fluid Intelligence, and Visual Processing Intelligence

Relatively few studies have examined the relationship between working memory and intelligence in children and, to our knowledge, none have evaluated working memory factors in relation to both *Gf* and *Gv*. Our results are consistent with previous research indicating that intelligence and working memory are related, but separable factors in children (Engel de Abreu, Conway, & Gathercole, 2010; Hornung et al., 2011).

By including *Gf*, whose relationship with working memory is thought to be driven by short-term storage (e.g. Colom, Abad, Quiroga, Shih and Flores-Mendoza. 2008; Hornung et al., 2011) or by attention or cognitive control (Cowan et al., 2006; Engel de Abreu, Conway, & Gathercole, 2010), and *Gv*, whose relationship should be driven by visuospatial working memory, we could evaluate the different contributions of the focus of attention and visuospatial processing to intelligence. Neither the Cowan nor Baddeley and Hitch models demonstrated a significant relationship between the central executive or phonological storage and rehearsal factors and *Gf* or *Gv*, yet both factors included tasks that required storage. This suggests that rather than storage driving the relationship with *Gf*, the focus of attention, as represented in the Cowan model, is the more likely candidate because when the relationship between the focus of attention (Cowan) or the visuospatial sketchpad factor (Baddeley and Hitch) with *Gv* was in the model, the focus of attention factor was a significant predictor of *Gf*, but the visuospatial sketchpad factor was not. In fact, in the Baddeley and Hitch model the only significant predictor of intelligence was the visuospatial sketchpad factor predicting *Gv* – none of the working memory factors predicted *Gf*. In contrast, in the Cowan working memory model the focus of attention factor was a significant predictor of both *Gf* (.52) and *Gv* (.76), as would be expected given the nature of the tasks, which included both visual and verbal stimuli.

One reason why the focus of attention may be highly related to intelligence is that it appears to be helpful for new concepts to form. Cowan, Donnell, and Saults (2013) demonstrated that items in the focus of attention together form new associations. They did this by presenting lists of 3, 6, or 9 words with an orienting task to determine which word was most interesting, and found that after this task ended, incidental long-term memory for which words had occurred in the same list were formed only for 3-word lists, i.e., lists short enough to be held in the focus of attention when presented initially. Halford, Cowan, and Andrews (2007) summarized evidence that the complexity of concepts that can be understood depends

on how many factors can be interrelated in working memory (presumably in the focus of attention). For example, the concept of a tiger depends on retaining concurrently the information that it is a type of cat (hence not a zebra) that is large (hence not a house cat) and striped (hence not a lion). Working memory limitations can help to explain children's common mistakes in word meaning.

Limitations

One limitation of our approach was our need to classify tasks a priori as likely indicators for a specific factor to use confirmatory factor analyses. In particular, it is not clear if attention, as a processing device, can be completely combined with attention as a storage device, as we have done out of practical necessity. The central executive component of working memory had a storage faculty in the conception of Baddeley and Hitch (1974), but not in later conceptions (e.g., Baddeley, 1986). Cowan et al. (2006) found attention-based storage and processing to be overlapping, but separable in developmental comparisons and regression analyses.

As a second related issue, although we were able to discern which tasks loaded on which factors, we do not have enough information to interpret unequivocally the differences in the strengths of loadings. For example, Cowan's central executive factor had a .77 loading onto n-back visual performance, much higher than any other variable. It may be that this particular task taps into both non-phonological storage and updating processes as mechanisms that require attention and central executive processes.

Concluding Remarks

The theoretical models of Baddeley and Hitch (1974; Baddeley, 1986), Cowan (1988, 1999, 2005), and Baddeley (2000) have inspired a considerable amount of research. They have done so not by making very specific predictions, but by putting forward grand schemes that are simple enough for various researchers to turn them over in the mind and have their own thoughts and intuitions about these models. For this research to be maximally effective, though, it is helpful to determine empirically the most important attributes that they must share and the most important attributes that distinguish them. Normative, experimental manipulations are diagnostic in that regard but, in a complementary way, so are studies of individual differences such as the present one. We show that any model of working memory would do well to have a component that is used for visual materials, for unpredictable verbal materials, and for binding between features; a component that differs between children and looks to us like the focus of attention (Cowan, 1988; Oberauer, 2002). The successful model also would do well to have a component that handles phonological materials of a predictable length, which also differs among children and may incorporate covert verbal rehearsal (although see Lewandowsky & Oberauer, 2015 for concerns). Finally, the model should include central executive processes that come into play more with materials that require manipulation than with those that need only to be retained. There are important new parietal brain correlates of the focus of attention (e.g., Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Majerus et al., 2016), frontal correlates of central executive storage-plus-manipulation (D'Esposito & Postle, 2015; Postle et al., 2006), and temporal and frontal correlates of storage and rehearsal of verbal materials (Buchsbaum, Olsen, Koch, & Berman, 2005). We

look forward to a new era in which researchers who frequent different journals, study different levels of analysis, use different methods, and come from different theoretical orientations can all contribute what they have learned to produce a general, robust model of the essential cognitive properties of the human mind.

Acknowledgments

References

Allen RJ, Baddeley AD, Hitch GJ. Is the binding of visual features in working memory resource-demanding? Journal of Experiment Psychology: General. 2006; 135:298–313.

Alloway TP. Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. European Journal of Psychological Assessment. 2009; 25:1–7.

Alloway TP, Cockcroft K. Working memory in ADHD: A comparison of British and South African children. Journal of Attention Disorders. 2014; 18:286–293. [PubMed: 22323119]

Alloway TP, Gathercole SE, Kirkwood H, Elliott J. The cognitive and behavioral characteristics of children with low working memory. Child Development. 2009; 80:606–621. [PubMed: 19467014]

Alloway TP, Gathercole SE, Pickering SJ. The automated working memory assessment. 2004 Test battery available from authors.

Alloway TP, Gathercole SE, Pickering SJ. Verbal and visuo-spatial short-term memory and working memory in children: Are they separable? Child Development. 2006; 77:1698–1716. [PubMed: 17107455]

Ang SY, Lee K. Central executive involvement in children's spatial memory. Memory. 2008; 16:918–933. [PubMed: 18802804]

Ang SY, Lee K. Exploring developmental differences in visual short-term memory and working memory. Developmental Psychology. 2010; 46:279–285. [PubMed: 20053024]

Baddeley, AD. Working memory. Oxford: Clarendon Press; 1986. Oxford Psychology Series #11

Baddeley A. The episodic buffer: a new component of working memory? Trends in Cognitive Sciences. 2000; 4:417–423. [PubMed: 11058819]

Baddeley, AD. Working memory, thought, and action. Oxford: Oxford University Press; 2007.

Baddeley, AD.; Hitch, GJ. Working memory. In: Bower, GA., editor. Recent advances in learning and motivation. Vol. 8. New York: Academic Press; 1974. p. 47-90.

Baddeley, AD.; Logie, RH. Working memory: The multiple component model. In: Miyake, A.; Shah, P., editors. Models of Working Memory: Mechanisms of active maintenance and executive control. Cambridge, U.K: Cambridge University Press; 1999. p. 28-61.

Bayliss DM, Jarrold C, Gunn MD, Baddeley AD. The complexities of complex span: Explaining individual differences in working memory in children and adults. Journal of Experimental Psychology: General. 2003; 32:71–92.

Bentler PM. Comparative fit indexes in structural models. Psychological Bulletin. 1990; 107:238–246. [PubMed: 2320703]

Blom E, Kuntay AC, Messer M, Verhagen J, Leseman P. The benefits of being bilingual: Working memory in bilingual Turkish-Dutch children. Journal of Experimental Child Psychology. 2014; 128:105–119. [PubMed: 25160938]

Broadway JM, Engle RW. Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. Behavior Research Methods. 2010; 42:563–570. [PubMed: 20479188]

Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. Testing structural equation models. Newbury Park, CA: Sage; 1993. p. 136-162.

Buchsbaum BR, Olsen RK, Koch P, Berman KF. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron. 2005; 48:687–97. [PubMed: 16301183]

Bunting MF, Cowan N, Colflesh GH. The deployment of attention in short-term memory tasks: Tradeoffs between immediate and delayed deployment. Memory & Cognition. 2008; 36:799–812. [PubMed: 18604962]

Bunting M, Cowan N, Saults JS. How does running memory span work? The Quarterly Journal of Experimental Psychology. 2006; 59:1691–1700. [PubMed: 16945854]

Cabbage KL, Brinkley S, Gray S, Alt M, Cowan N, Green S, Kuo T, Hogan TP. The comprehensive assessment battery for children – working memory (CABC-WM). 2016 Manuscript submitted for publication.

Camos V, Barrouillet P. Developmental change in working memory strategies: From passive maintenance to active refreshing. Developmental Psychology. 2011; 47:898–904. [PubMed: 21417518]

Camos V, Mora G, Oberauer K. Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. Memory & Cognition. 2011; 39:231–244. [PubMed: 21264630]

Campos IS, Almeida LS, Ferreira AI, Martinez LF. Working memory as separable subsystems: A study with Portuguese primary school children. Spanish Journal of Psychology. 2013; 16:1–10.

Carroll, JB. Human cognitive abilities: A survey of factor-analytic studies. New York, NY: Cambridge University Press; 1993.

Colom R, Abad FJ, Quiroga MA, Shih PC, Flores-Mendoza C. Working memory and intelligence are highly related constructs, but why? Intelligence. 2008; 36:584–606.

Conway ARA, Cowan N, Bunting MF, Therriault DJ, Minkoff SRB. A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. Intelligence. 2002; 30:163–183.

Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. Psychological Bulletin. 1988; 104:163–191. [PubMed: 3054993]

Cowan, N. Attention and memory: An integrated framework. Oxford, UK: Oxford University Press; 1995a.

Cowan N. Verbal working memory: A view with a room. American Journal of Psychology. 1995b; 108:123–155.

Cowan, N. An embedded-processes model of working memory. In: Miyake, A.; Shah, P., editors. Model of working memory: Mechanism of active maintenance and executive control. Cambridge, U.K: Cambridge University Press; 1999. p. 62-101.

Cowan N. The magical number four in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences. 2001; 24:87–114. [PubMed: 11515286]

Cowan, N. Working memory capacity. Hove, East Sussex, UK: Psychology Press; 2005.

Cowan N, Donnell K, Saults JS. A list-length constraint on incidental item-to-item associations. Psychonomic Bulletin & Review. 2013; 20:1253–1258. [PubMed: 23653420]

Cowan N, Elliott EM, Saults JS, Morey CC, Mattox S, Hismjatullina A, Conway ARA. On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. Cognitive Psychology. 2005; 51:42–100. [PubMed: 16039935]

Cowan N, Fristoe NM, Elliott EM, Brunner RP, Saults JS. Scope of attention, control of attention, and intelligence in children and adults. Memory & Cognition. 2006; 34:1754–1768. [PubMed: 17489300]

Cowan N, Saults JS, Blume CL. Central and peripheral components of working memory storage. Journal of Experimental Psychology: General. 2014; 143:1806–1836. [PubMed: 24867488]

Crowder RG. Behavioral strategies in immediate memory. Journal of Verbal Learning and Verbal Behavior. 1969; 8:524–528.

D'Esposito M, Postle BR. The cognitive neuroscience of working memory. Annual Review of Psychology. 2015; 66:115–42.

Elosúa MR, Ruiz RM. Absence of hardly pursued updating in a running memory task. Psychological Research. 2008; 123:451–460.

Engel de Abreu PMJ, Conway ARA, Gathercole SE. Working memory and fluid intelligence in young children. Intelligence. 2010; 38:552–561.

Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General. 1999; 128:309–331. [PubMed: 10513398]

Flavell JH, Beach DH, Chinsky JM. Spontaneous verbal rehearsal in a memory task as a function of age. Child Development. 1966; 37:283–299. [PubMed: 5941895]

Gabig CS. Verbal working memory and story retelling in school-age children with autism. Language, Speech, and Hearing Services in Schools. 2008; 39:498–511.

Gignac GE, Watkins MW. There may be nothing special about the association between working memory capacity and fluid intelligence. Intelligence. 2015; 52:18–23.

Goldman, R.; Fristoe, M. Goldman-Fristoe Test of Articulation 2. Circle Pines, MN: American Guidance Service, Inc; 2000.

Gray S. The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. Journal of Speech, Language, and Hearing Research. 2006; 49:955–969.

Gray S, Alt M, Hogan TP, Green S, Cowan N. Comprehensive Assessment Battery for Children Working Memory. nd Unpublished measure under development.

Green SB, Yang Y. Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficient. Educational Measurement: Issues and Practice. 2015

Green SB, Yang Y, Alt M, Brinkley S, Gray S, Hogan T, Cowan N. Use of internal consistency coefficients for estimating reliability of experimental task scores. Psychonomic Bulletin & Review. in press.

Guttentag RE. The mental effort requirement of cumulative rehearsal: A developmental study. Journal of Experimental Child Psychology. 1984; 37:92–106.

Halford GS, Cowan N, Andrews G. Separating cognitive capacity from knowledge: A new hypothesis. Trends in Cognitive Sciences. 2007; 11:236–242. [PubMed: 17475538]

Hockey R. Rate of presentation in running memory and direct manipulation of input processing strategies. Quarterly Journal of Experimental Psychology. 1973; 25:104–111.

Hornung C, Brunner M, Reuter RAP, Martin R. Children's working memory: Its structure and relationship to fluid intelligence. Intelligence. 2011; 39:210–221.

Hu Y, Allen RJ, Baddeley AD, Hitch GJ. Executive control of stimulus-driven and goal-directed attention in visual working memory. Attention, Perception, & Psychophysics. 2016; E-pub ahead of print. doi: 10.3758/s13414-016-1106-7

Jeffries S, Everatt J. Working memory: Its role in dyslexia and other specific learning disabilities. Dyslexia. 2004; 10:196–214. [PubMed: 15341198]

Kane MJ, Engle RW. Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. Journal of Experimental Psychology: General. 2003; 132:47–70. [PubMed: 12656297]

Kaufman, AS.; Kaufman, NL. Kaufman Assessment Battery for Children. 2. Circle Pines, MN: American Guidance Service, Inc; 2004.

Kuhn, J. Controlled attention and storage: An investigation of the relationship between working memory, short-term memory, scope of attention, and intelligence in children. Learning and Individual Differences. in press http://dx.doi.org/10.1016/j.lindif.2015.04.009

Leonard LB, Ellis Weismer S, Miller CA, Francis DJ, Tomblin JB, Kail RV. Speed of processing, working memory, and language impairment in children. Journal of Speech, Language, and Hearing Research. 2007; 50:408–428.

Lewandowsky S, Oberauer K. Rehearsal in serial recall: An unworkable solution to the nonexistent problem of decay. Psychological Review. 2015; 122:674–99. [PubMed: 26437148]

Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. Neural evidence for a distinction between short-term memory and the focus of attention. Journal of Cognitive Neuroscience. 2012; 24:61–79. [PubMed: 21955164]

Mackintosh NJ, Bennett ES. The fractionation of working memory maps onto different components of intelligence. Intelligence. 2003; 31:519–531.

Majerus S, Cowan N, Péters F, Van Calster L, Phillips C, Schrouff J. Cross-modal decoding of neural patterns associated with working memory: Evidence for attention-based accounts of working memory. Cerebral Cortex. 2016; 26:166–179. [PubMed: 25146374]

McGrew, KS. The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In: Flanagan, DP.; Genshaft, JL.; Harrison, PL., editors. Contemporary intellectual assessment: Theories, tests, and issues. New York: Guilford; 2005.

McGrew KS, Wendling BJ. Cattell–Horn–Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. Psychology in the Schools. 2010; 47:651–675.

Michalczyk K, Malstadt N, Worgt M, Konen T, Hasselhorn M. Age differences and measurement invariance of working memory in 5- to 12-year-old children. European Journal of Psychological Assessment. 2013; 29:220–229.

Montgomery JW, Evans JL. Complex sentence comprehension and working memory in children with specific language impairment. Journal of Speech, Language, and Hearing Research. 2009; 52:269–288.

Morales J, Calvo A, Bialystok E. Working memory development in monolingual and bilingual children. Journal of Experimental Child Psychology. 2013; 114:187–202. [PubMed: 23059128]

Muthén, LK.; Muthén, BO. 2007 Mplus user's guide. 6. Los Angeles, CA: Muthén & Muthén; 1998–2012.

Nevo E, Breznitz Z. The development of working memory from kindergarten to first grade in children with different decoding skills. Journal of Experimental Child Psychology. 2013; 114:217–228. [PubMed: 23073369]

Oberauer K. Access to information in working memory: exploring the focus of attention. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28:411–421.

Oberauer K, Schultze R, Wilhelm O, Suß H. Working memory and intelligence–their correlation and their relations: Comment on Ackerman, Beier, and Boyle (2005). Psychological Bulletin. 2005; 131:61–65. [PubMed: 15631551]

Ornstein, PA.; Naus, MJ. Rehearsal processes in children's memory. In: Ornstein, PA., editor. Memory development in children. Hillsdale, NJ: Erlbaum; 1978. p. 69-99.

Paolozza A, Rasmussen C, Pei J, Hanlong-Dearman A, Nikkel SM, et al. Working memory and visuospatial deficits correlate with oculomotor control in children with fetal alcohol spectrum disorder. Behavioural Brain Research. 2014; 263:70–79. [PubMed: 24486257]

Pickering, SJ.; Gathercole, S. Working Memory Test Battery for Children: Manual. London, UK: Psychological Corporation; 2001.

Postle BR, Ferrarelli F, Hamidi M, Feredoes E, Massimini M, Peterson M, Alexander A, Tononi G. Repetitive transcranial magnetic stimulation dissociates working memory manipulation from retention functions in the prefrontal, but not posterior parietal, cortex. Journal of Cognitive Neuroscience. 2006; 18:1712–1722. [PubMed: 17014375]

Raju NS. A generalization of coefficient alpha. Psychometrika. 1977; 42:549–565.

Raven, J.; Raven, JC.; Court, JH. Manual for the Raven's Progressive Matrices and Vocabulary Scales (1998 Edition). Oxford, UK: Oxford Psychologists Press; 1998. Section 2: Coloured Progressive matrices.

Reynolds MR, Keith TZ, Flanagan DP, Alfonso VC. A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. Journal of School Psychology. 2013; 51:535–555. [PubMed: 23870446]

Semel, E.; Wiig, EH.; Secord, WA. Clinical Evaluation of Language Fundamentals–Fourth Edition. Circle Pines, MN: American Guidance Service, Inc; 2003.

Shah P, Miyake A. The separability of working memory resources for spatial thinking and language processing: An individual differences approach. Journal of Experimental Psychology: General. 1996; 125:4–27. [PubMed: 8851737]

Shahabi SR, Abad FJ, Colom R. Short-term storage is a stable predictor of fluid intelligence whereas working memory capacity and executive function are not: A comprehensive study with Iranian schoolchildren. Intelligence. 2014; 44:134–141.

Swanson HL. Intellectual growth in children as a function of domain specific and domain general working memory subgroups. Intelligence. 2011; 39:481–492.

Tillman CM, Bohlin G, Sorensen L, Lundervold AJ. Intelligence and specific cognitive abilities in children. Journal of Individual Differences. 2009; 30:209–219.

Torgesen, J.; Wagner, R.; Rashotte, C. Test of Word Reading Efficiency. 2. Circle Pines, MN: American Guidance Service, Inc; 2012.

Unsworth N, Engle RW. On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. Psychological Bulletin. 2007; 133:1038–1066. [PubMed: 17967093]

Van der Molen MJ. Working memory structure in 10- and 15-year old children with mild to borderline intellectual disabilities. Research in Developmental Disabilities. 2010; 31:1258–1263. [PubMed: 20832241]

Van der Molen MJ, Henry LA, Van Luit JEH. Working memory development in children with mild to borderline intellectual disabilities. Journal of Intellectual Disability Research. 2014; 58:637–650. [PubMed: 23802604]

Appendix

Description of Experimental Tasks included in the Comprehensive Assessment Battery for Children – Working Memory (CABC-WM; Gray, Alt, Hogan, Green, & Cowan, Under Development)

| Task	Stimuli	Trial Types	Number of Training Blocks and Trials (in parentheses)	Number of Trials Correct to Pass Training	Number of Trials and Stimuli	Task Length (min)	Dependent Variable(s)
Central Executive Tasks							
N-back Auditory	• Image of robot band • Tones	• •	1 training block: Same • Different •	4/6 Same (3) Different (3)	54 (3 blocks each with 9 Same, 9 Different)	6.50	Mean accuracy for same and different trials combined
N-back Visual	• Images of black squares with white dots	• •	1 training block: Same • Different •	4/6 Same (3) Different (3)	54 (3 blocks each with 9 Same, 9 Different)	7.50	Mean accuracy for same and different trials combined
Number Updating	• Visual presentation of numbers and operations	Not applicable	2 training blocks: • Each block (5)	5/5 each block	15 (3 blocks each with 5 trials)	7.20	Mean accuracy for all trials
Short-Term Phonological Memory Tasks							
Digit Span	• Auditory recordings of digits 1–9 (except 7 because it is 2 syllables)	Span length (2–8 digits)	1 training block: • (2)	2/2	14 (2 trials at each span length from 2–8 digits)	4.50	Number of trials correct at each span length x span length then sum products
Digit Span – Running	• Auditory recordings of digits 1–9 (except 7 because it is 2 syllables)	Span length (7–10 digits)	3 training blocks: • Each block (3)	At least 1 correct for each of 3 blocks	12 (3 trials at each span length from 7–10 digits)	6.00	Average number of digits recalled in the correct order
Nonword Repetition	• Auditory recordings of nonwords	Word length (2–5 syllable nonwords)	1 training block: • (3 2-syllable trials)	3 attempted	16 nonwords (4 each at 2-, 3-, 4- and 5-	3.00	Number of words repeated with correct consonants at each syllable length x syllable length then sum products

Task	Stimuli	Trial Types	Number of Training Blocks and Trials (in parentheses)	Number of Trials Correct to Pass Training	Number of Trials and Stimuli	Task Length (min)	Dependent Variable(s)
					syllable lengths)		
Short-Term Visuospatial Memory Tasks							
Location Span	• An arrow pointing toward a location arranged in a circular pattern	Span length (2–6 locations)	3 training blocks: • 1 location (1) • 2 locations (2)	At least 1 at 1 location and 1 at 2 locations	12 (2 trials at each span length from 2–6 locations)	4.50	Correct number of trials at each span length x span length then sum products
Location Span Running	• An arrow pointing toward a location arranged in a circular pattern	Span length (5–8 locations)	3 training blocks: • 6 locations (1) • 7 locations (1) • 8 locations (1)	1/1 correct at each length	12 (3 trials at each span length from 5–8 locations)	7.50	Average number of locations correctly identified across all trials
Visual Span	• Black polygons	Span length (1–6 polygons)	1 training block: • 1 polygon (1) • 2 polygons (2)	3/3	12 (2 trials at each span length from 1–6)	6.50	Correct number of trials at each span length x span length then sum products
Visual Span – Running	• Black polygons	Span length (3–6 polygons)	1 training block: • 3 polygons (1) • 4 polygons (1)	1 correct at each length	12 (3 trials at each span length from 3–6 polygons)	7.00	Average number of polygons correctly identified in order across all trials
Binding Tasks							

Task	Stimuli	Trial Types	Number of Training Blocks and Trials (in parentheses)	Number of Trials Correct to Pass Training	Number of Trials and Stimuli	Task Length (min)	Dependent Variable(s)
Phonological Binding Span	• Auditory non-speech sounds (e.g. mechanical noises) • Auditory recordings of nonwords	Span Length (1–4)	1 training block: • 1 sound-nonword pair (1) • 2 sound-nonword pairs (1)	Attempt 2/2	20 sound-nonword pairs (2 trials each of 1–4 pairs per trial)	5.20	Correct number of trials at each span length x span length then sum products
Visual-Spatial Binding Span	• Image of a grid • Black polygons	Span Length (1–6 polygons)	1 training block: • 1 polygon (1) • 2 polygons (1)	2/2	12 (2 trials at each span length from 1–6 polygons)	5.20	Correct number of trials at each span length x span length then sum products
Cross-Modal Binding	• Black polygons • Auditory recordings of nonwords	Span Length (1–6 polygons)	1 training block: 1 nonword-polygon pair (1) 2 nonword-polygon pairs (1)	2/2	12 (2 trials at each span length from 1–6 polygons)	6.50	Correct number of trials at each span length x span length then sum products

Highlights

- Tested the fit of three competing theoretical models of working memory in children.

- Results suggest that working memory models are coming together on common ground.

- No evidence for a separate episodic buffer factor (Baddeley, 2000)

- Focus of attention predicted fluid reasoning and visual processing intelligence.
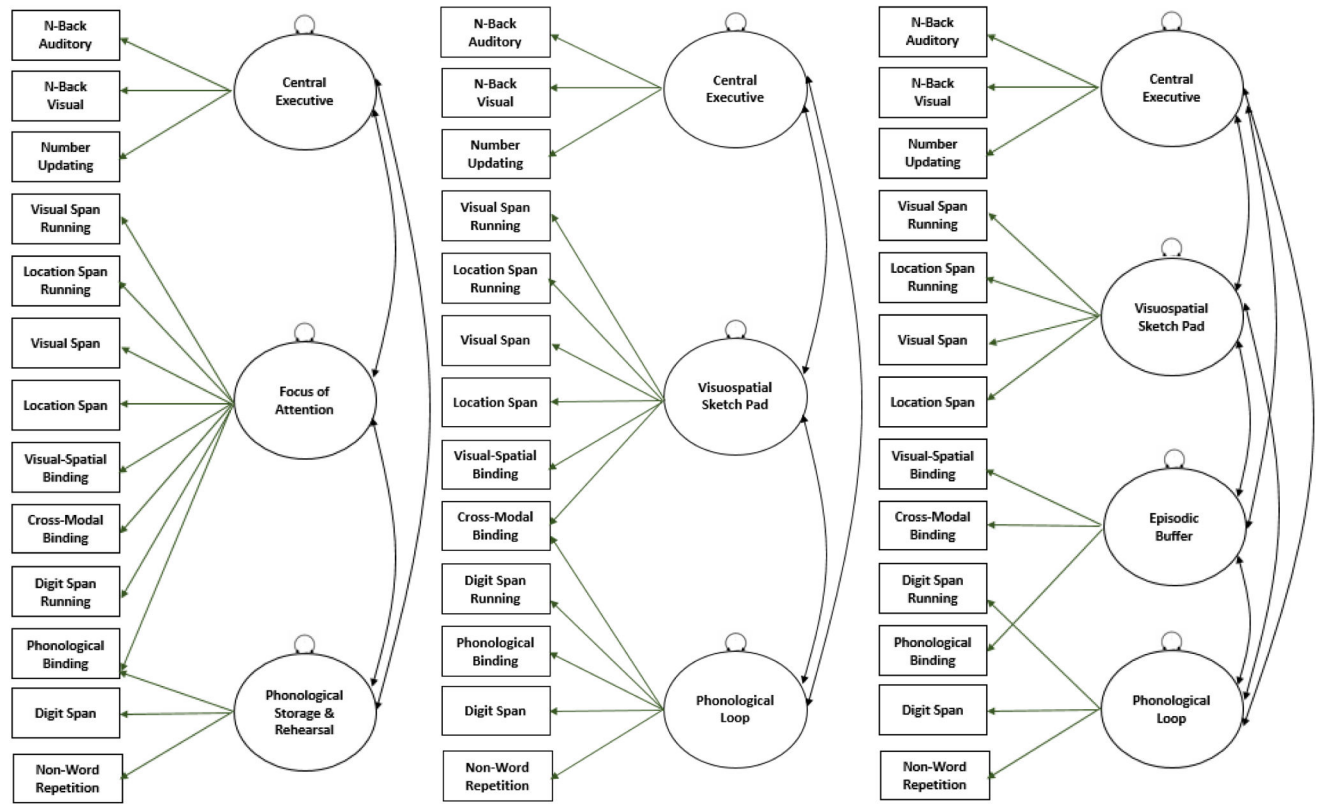
Cowan's Embedded Processes Model

Baddeley and Hitch's (1974) Three-Component Model

Baddeley's (2000) Four-Component Model

Figure 1.
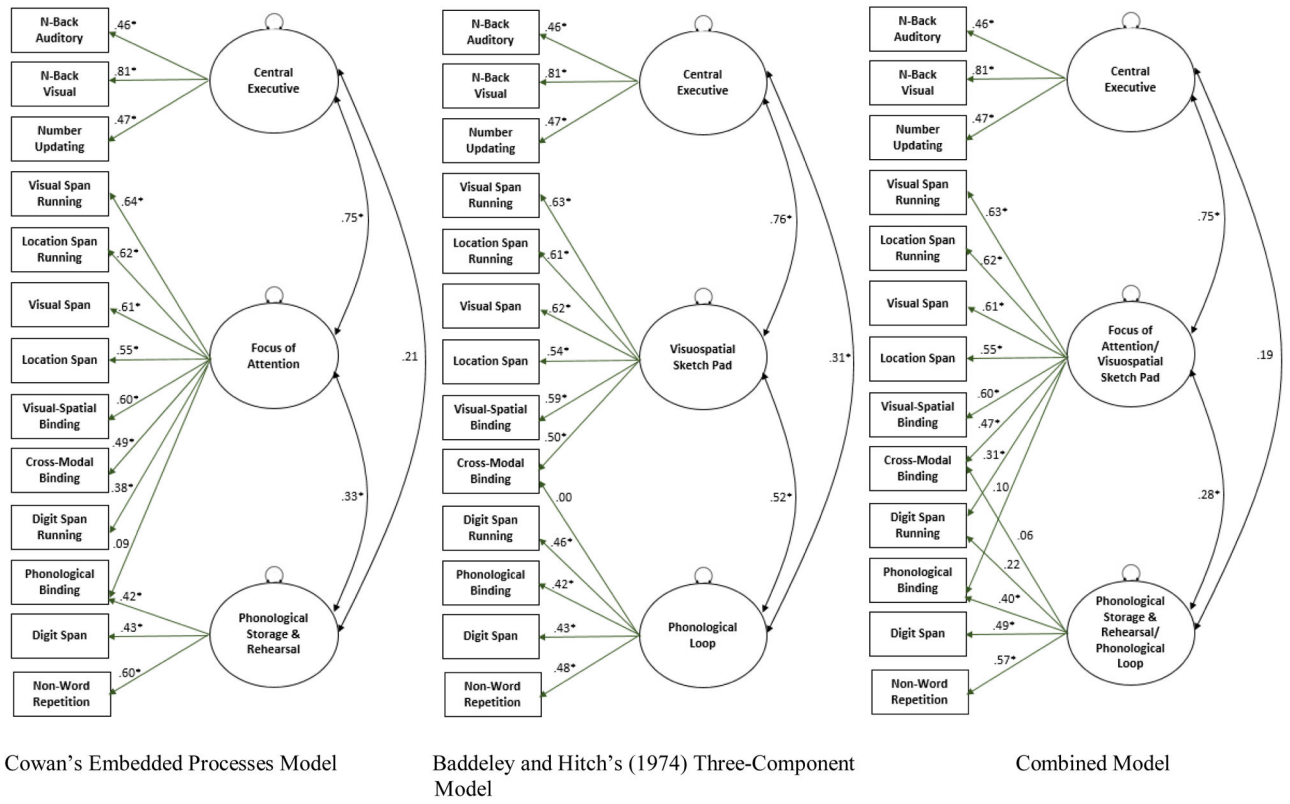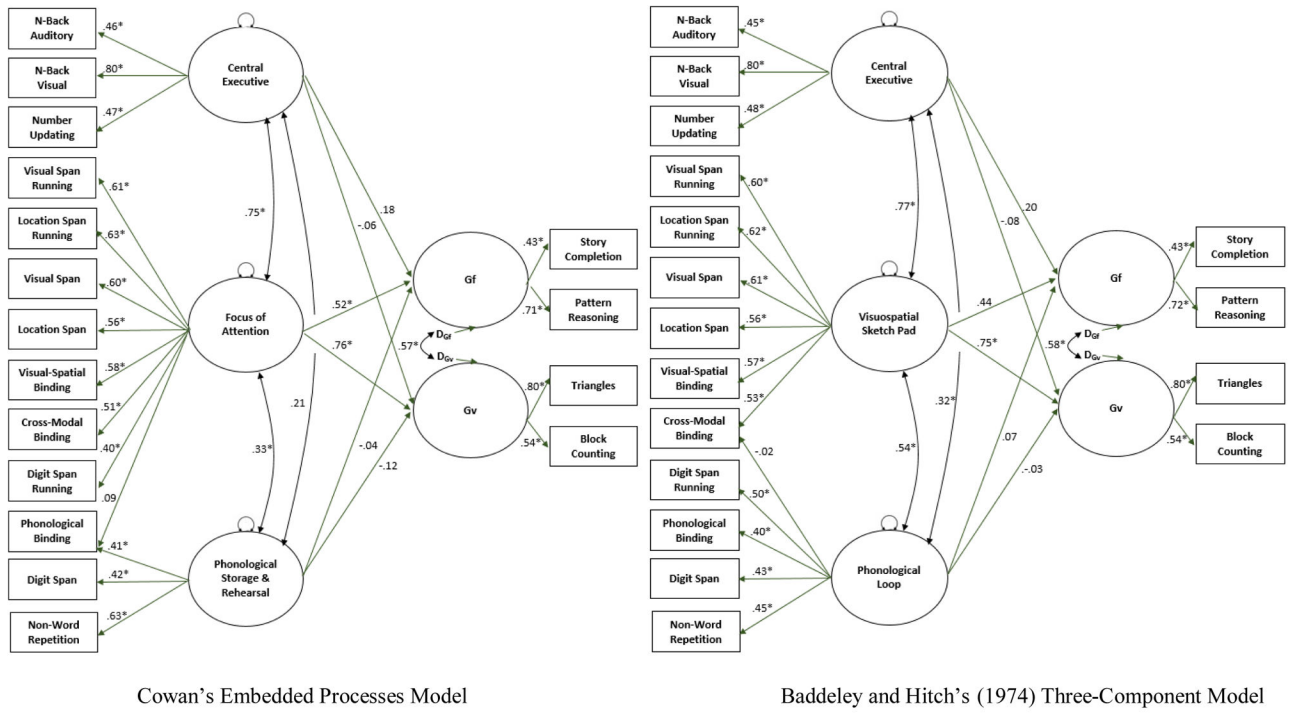Hypothesized working memory models.

Figure 2.
Standardized coefficients for three working memory models.

Cowan's Embedded Processes Model   Baddeley and Hitch's (1974) Three-Component Model

Figure 3.
Standardized coefficients for working memory models with intelligence factors.

Table 1

Summary of Studies Investigating the Structure of Working Memory in Children with Typical Development

Authors	Ages N Language	Working Memory Components Assessed (Test Name)	Number of Indicators per Component	Factors in Final Model
Bayliss, Jarrold, Gunn, & Baddeley (2003)	7–9 yrs N=75 English	Complex span	4	General processing
		Storage	2	Verbal storage
		Processing	2	Visuospatial storage
Alloway, Gathercole Willis, & Adams (2004)	4–6 yrs N = 633 English	Verbal short-term memory	2	Central executive
		Complex memory span	3	Episodic buffer
		Episodic buffer (WMTB-C plus other tasks)	2	Phonological loop
Alloway, Gathercole, & Pickering (2006)	4–11 yrs N = 709 English	Verbal short-term memory	3	Domain general mechanism
		Verbal working memory	3	Verbal storage
		Visuospatial short-term memory	3	Visuospatial storage
		Visuospatial working memory (AWMA)	3	
Hornung, Brunner Reuter, & Martin (2011)	5–8 yrs N = 161 Luxemburgish or Portuguese	Verbal simple span	2	Central executive attention
		Verbal complex span	2	Verbal
		Visuo-spatial span (WMTB-C)	2	Visuospatial
Campos, Almeida, Ferreira, & Martinez (2013)	7–9 yrs N = 103 Portuguese	Central executive function	3	Central executive + visuospatial sketchpad
		Phonological loop function	4	Phonological loop
		Visuospatial sketchpad function (WMTB-C)	2	
Giofre, Mammarella, & Cornoldi (2013)	8–10 yrs N = 176 Italian	Simple span	3	Attentional control system
		Matrices span (visuospatial)	2	Domain specific verbal
		Categorization working memory span	1	Domain specific visuospatial
		Listening span	1	
		Visual pattern test, active	1	
Michalczyk, Malstadt, Worgt, Konen, & Hasselhorn (2013)	5–6 (284) 7–9 (690)	Central executive/inhibition	6	Central executive
		Phonological loop	4	Phonological loop
	10–12 (695) N = 1,669 German	Visuospatial sketchpad (WMTB-C, German version)	2	Visuospatial sketchpad
Nadler & Archibald (2014)	5–9 N=178 English	Phonological short-term memory	3	Visuospatial short-term memory
		Verbal working memory	3	Phonological short-term memory

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Authors	Ages N Language	Working Memory Components Assessed (Test Name)	Number of Indicators per Component	Factors in Final Model
		Visuospatial short-term memory	3	Central executive
		Visuospatial working memory (AWMA)	3	
Gray, Green, Alt, Hogan, Kuo, Brinkley, & Cowan (present study)	7–9 N = 168 English	Central executive	3	See results section
		Phonological loop	3	
		Visuospatial sketchpad	4	
		Episodic buffer (ABC-WM)	3	

Note. WMTB-C = Working Memory Test Battery for Children (Pickering & Gathercole, 2001); AWMA = Automated Working Memory Assessment battery (Alloway, Gathercole, & Pickering, 2004); ABC-WM = Assessment Battery for Children – Working Memory (Gray, Alt, Hogan, Green, & Cowan, under development).

Table 2

Participant Description Information Including Summary of Inclusionary Test Results (N=168)

Measure	M	SD	Range
Age (years; months)	7; 9	0; 5	7;0 – 9;1
Mother's Education in Years	15.40	1.67	12 – 17
GFTA-2 Articulation Accuracy Percentile	50.65	9.01	7 – 62
K-ABC II Nonverbal Index Standard Score	117.60	15.48	78 – 160
TOWRE-2 Word/Nonword Standard Score	109.39	8.43	96 – 145
CELF-4 Core Language Standard Score	108.71	9.57	88 – 130

Note. GFTA-2 = Goldman-Fristoe Test of Articulation –2 (Goldman & Fristoe, 2000); K-ABC II = Kaufman Assessment Battery for Children, Second Edition (Kaufman & Kaufman, 2004); TOWRE-2 = Test of Word Reading Efficiency, Second Edition (Torgesen, Wagner, & Rashotte, 2012); CELF-4 = Clinical Evaluation of Language Fundamentals – Fourth Edition (Semel, Wiig, & Secord, 2003).

Table 3

Reliability of Working Memory Tasks

Type of Task (type of score)	*N*	Reliability	95% CI
Number Updating (accuracy)	139	.95	[.93, .96]
N-back Visual (accuracy)	148	.86	[.81, .90]
N-back Auditory (accuracy)	151	.82	[.75, .87]
Location Span (weighted sum)	158	.70	[.59, .78]
Location Span running (mean)	146	.93	[.91, .95]
Visual Span (weighted sum)	140	.73	[.62, .81]
Visual Span Running (mean)	99	.84	[.78, .89]
Digit Span (weighted sum)	159	.67	[.55, .76]
Digit Span Running (mean)	109	.85	[.79, .89]
Nonword Repetition (weighted sum)	153	.60	[.45, .71]
Cross-Modal Binding (mean)	153	.38	[.15, .55]
Phonological Binding Span (weighted	149	.53	[.35, .66]
Visual-Spatial Binding (weighted sum) sum)	145	.51	[.32, .65]

Table 4

Correlations[a] among Task and KABC Measures (Lower-Left Triangle), N for Each Measure (Diagonal), and N for Each Pair of Measures (Upper-Right Triangle)

Measures	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 N-back Auditory	155	147	153	154	153	149	153	153	154	127	146	143	151	155	155	155	154
2 N-back Visual	.37**	155	153	154	153	149	153	153	153	126	145	145	151	155	155	155	154
3 Number Updating	.26**	.34**	163	163	161	158	162	162	159	133	154	148	159	163	163	163	162
4 Digit Span	.01	.07	.00	164	162	159	163	163	160	134	155	149	160	164	164	164	163
5 Digit Span Running	.06	.19*	.09	.23**	163	158	162	161	159	134	154	148	161	163	163	163	162
6 Nonword Repetition	.08	.18*	.01	.27**	.16*	159	159	158	155	131	153	144	156	159	159	159	158
7 Location Span	.13	.27**	.21**	.11	.22**	.04	164	162	159	134	156	148	160	164	164	164	163
8 Location Span Running	.06	.41**	.28**	.07	.29**	.08	.43**	163	159	133	154	148	159	163	163	163	162
9 Visual Span	.24**	.45**	.27**	.05	.17*	.17*	.37**	.32**	161	131	151	148	157	161	161	161	160
10 Visual Span Running	.14	.38**	.22*	.09	.32**	.22*	.26**	.40**	.32**	134	129	121	132	134	134	134	133
11 Phonological Binding	.07	.04	.18*	.21**	.10	.27**	.10	.18*	.13	.22*	156	142	153	156	156	156	155
12 Visual-Spatial Binding	.10	.27**	.19*	-.08	.22**	.05	.33**	.36**	.39**	.46**	.07	150	146	150	150	150	149
13 Cross-Modal Binding	.22**	.33**	.23**	.09	.12	.13	.27**	.22**	.31**	.40**	.16	.28**	161	161	161	161	160
14 KABC-Block Counting	-.09	.17*	.18*	.04	.29**	.09	.18*	.26**	.17*	.20*	.12	.18*	.19*	167	167	167	166
15 KABC-Triangles	.19*	.34**	.17*	.16*	.19*	-.02	.37**	.39**	.32**	.26**	.07	.22**	.33**	.43**	167	167	166
16 KABC-Pattern Reasoning	.17*	.32**	.26**	.10	.29**	.09	.24**	.27**	.30**	.26**	.09	.26**	.32**	.32**	.40**	167	166
17 KABC-Story Completion	.09	.20*	.08	-.06	-.01	-.01	.15	.15	.21**	-.04	.01	.04	.26**	.16*	.31**	.31**	166
Mean	.84	.77	.80	19.58	1.85	11.47	10.77	1.33	6.98	.88	12.17	4.75	4.35	12.11	12.22	13.18	12.41
Standard deviation	.14	.17	.27	6.84	1.24	6.63	6.11	.66	5.65	.65	6.93	3.17	2.65	2.32	2.93	3.07	2.69

[a] The correlations are computed using pairwise deletion. The results in the paper were not based on the statistics reported in the table, but were computed based on raw data and maximum likelihood estimation with robust standard errors (MLR).

[*] $p < .05$,

[**] $p < .01$

Table 5

Fit for Working Memory Models and Models Examining Relationships between Working Memory and Intelligence Factors

Model	χ^2 test	RMSEA 95% CI	CFI	AIC
Working Memory Models				
Cowan Embedded Process	$\chi^2(61) = 70.83$, $p = .183$	.031 [.000, .059]	.968	7436
Baddeley-Hitch	$\chi^2(61) = 77.90$, $p = .071$	.041 [.000, .065]	.945	7440
Combined	$\chi^2(61) = 67.69$, $p = .205$	.030 [.000, .058]	.971	7436
Working Memory—Intelligence Models				
Cowan Embedded Process	$\chi^2(108) = 137.29$, $p = .030$	.040 [.013, .059]	.937	10552
Baddeley-Hitch	$\chi^2(108) = 147.23$, $p = .007$	.047 [.025, .065]	.916	10558

Table 6

Correlations between Working Memory Factors and Intelligence Factors

Working memory factors	Gf	GV
Cowan embedded process model		
Central executive	.57*	.49*
Focus of attention	.64*	.68*
Phonological storage & rehearsal	.16	.11
Baddeley-Hitch model		
Central executive	.56*	.49*
Visuospatial sketch pad	.63*	.68*
Phonological loop	.37	.35

p < .05