Working memory capacity and the scope and control of attention

Zach Shipstead[1] · Tyler L. Harrison[2] · Randall W. Engle[2]

Abstract Complex span and visual arrays are two common measures of working memory capacity that are respectively treated as measures of attention control and storage capacity. A recent analysis of these tasks concluded that (1) complex span performance has a relatively stronger relationship to fluid intelligence and (2) this is due to the requirement that people engage control processes while performing this task. The present study examines the validity of these conclusions by examining two large data sets that include a more diverse set of visual arrays tasks and several measures of attention control. We conclude that complex span and visual arrays account for similar amounts of variance in fluid intelligence. The disparity relative to the earlier analysis is attributed to the present study involving a more complete measure of the latent ability underlying the performance of visual arrays. Moreover, we find that both types of working memory task have strong relationships to attention control. This indicates that the ability to engage attention in a controlled manner is a critical aspect of working memory capacity, regardless of the type of task that is used to measure this construct.

Keywords Working memory capacity · Maintenance · Storage · Attention control · Fluid intelligence

✉ Zach Shipstead
zach.shipstead@asu.edu

1 Department of Social and Behavioral Sciences, Arizona State University, Glendale, AZ 85306-4908, USA

2 School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332, USA

Introduction

Working memory is the cognitive system in which memory and attention interact to produce complex cognition. Working memory *capacity* is a measure of individual differences in the efficacy with which this system functions. These differences are important, as they predict performance on tests of academic achievement (Cowan et al., 2005; Turner & Engle, 1989), language comprehension (Daneman & Mirikle, 1996), and performance differences often remain despite extensive training and experience in a domain (Hambrick & Meinz, 2011).

Nelson Cowan and colleagues note that individual difference in working memory capacity are often ascribed to at least one of two broadly defined mechanisms: the scope or control of attention (Cowan et al., 2005; Cowan et al., 2006; see also Chuderski & Nęcka, 2012; Shipstead, Redick, Hicks, & Engle, 2012; Shipstead et al., 2014). The scope of attention refers to individual differences in the maintenance capacity of focal attention (Cowan, 2001). The benefit of a large maintenance capacity is that it allows a person to protect more information from proactive interference, thus reducing the tendency to lose sight of goals and increasing the likelihood that disparate units of information will be combined during novel reasoning. The control of attention refers to processes that ensure appropriate information is selected and stabilized in focal attention. This second mechanism increases the functional capacity of focal attention by ensuring that task-relevant, rather than irrelevant, information enters focal attention (see Awh & Vogel, 2008; McNab & Klingberg, 2008; Vogel, McCullough, & Machizawa, 2005). In the present study the control of attention will be operationally defined through tasks that require the override of prepotent responses, in favor of goal-relevant actions (antisaccade, Stroop and flanker tasks; see Methods). Thus, it is critical to state that we are presently referring to an ability to avoid being drawn into salient distraction.

As an example of how these mechanisms might function together, consider a fluid intelligence problem solving task (e.g., matrix reasoning, sequence completion). The test-taker is presented with several problem components and several possible solutions. In generating problem solving hypotheses, a large focus of attention will allow for broader integration of information, as more components can be considered at any one point in time (Oberauer et al., 2007). At the same time, broadly-considered-hypotheses are not necessarily correct hypotheses. A test-taker may need to alter a solution, consider information that is relevant to a hypothesis, or retrieve an old solution for further consideration. In these cases, the ability to use attention to resolve competition for retrieval into focal attention will become critical to task performance (e.g., using attention control to combat proactive interference generated by initial problem-solving attempts).

Previous examination of this distinction

Measuring the scope of focal attention

Shipstead, Redick et al. (2012) examined the scope and control of attention using two types of working memory capacity tests: visual arrays and complex span. The visual arrays task (Fig. 1a) is the quintessential measure of the maintenance capacity of working memory (Awh, Barton, & Vogel, 2007; Cowan, 2001; Cowan et al., 2005, 2006; Fukuda et al., 2010; Luck & Vogel, 1997; Morey & Cowan, 2004, 2005; Vogel, McCollough, & Machizawa, 2005). In this task an array of objects (colored boxes, line orientations, etc.) is momentarily displayed on a computer monitor, and then disappears. After a brief interval (~1 s), the array reappears and the test-taker simply responds as to whether or not one of the objects has changed (new color, new orientation).

The elegance of this task is that it does not involve any distraction or complex processing. A person simply maintains a memory of the array. In general people have near-perfect accuracy if the array contains 1–3 objects. With four objects
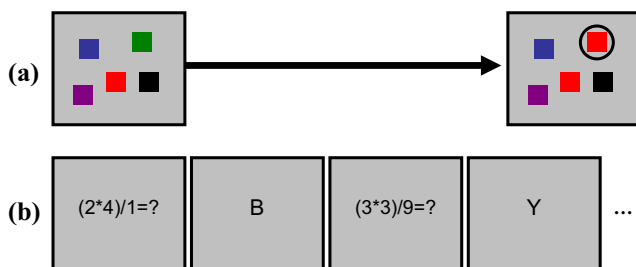
accuracy begins to decline. With each additional object, a steady decrease in accuracy is found (see Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001). This trend is interpreted as evidence that most people can maintain between 3–4 items in focal attention and, as more items are added to the display, the likelihood of guessing increases.

Measuring the control of the contents of focal attention

Complex span tasks are assumed to put a heavy burden on controlled processing and controlled maintenance during periods of distraction (Daneman & Merikle, 1996; Engle, 2002; Engle & Kane, 2004). Figure 1b displays the operation span task. In this task, a test-taker is presented with a simple mathematical equation that must be solved. After this is completed the test-taker is presented with a to-be-remembered letter, followed by a new equation. After 3–7 such equation/item pairs the test-taker is cued to recall all of the letters in the order that they were originally presented.

The challenge presented by complex span tasks is one of remembering the items, despite being distracted by the processing task. Thus, this task provides an ideal context for studying working memory capacity as the ability to control the contents of focal attention. The demands that complex span tasks place on working memory include deployment of attention to maintain to-be-remembered items when conscious processes are otherwise occupied (Kane et al., 2007; Shipstead et al., 2014) and retrieval of information from longer-term storage when attention fails (Unsworth & Engle, 2006a). Indeed, it has been repeatedly demonstrated that people with high scores on complex span tasks also show superior performance on measures of attention control (Conway, Bunting, & Cowan, 2001; Kane et al., 2001; Kane & Engle, 2003; Redick & Engle, 2006; Shipstead, Harrison, & Engle, 2012; Unsworth et al., 2004) and memory retrieval (Healy & Miyake, 2009; Shipstead et al., 2014; Unsworth, 2009; Unsworth et al., 2014; Unsworth & Spillers, 2010), relative to people with low scores.

Previous examination

From these perspectives, it is reasonable to assume that performance on visual arrays and complex span tasks reflects complimentary aspects of the overall working memory system. Visual arrays tasks reflect raw maintenance capacity of focal attention, while complex span tasks reflect abilities that ensure only appropriate information is maintained by, or recalled into, focal attention. With this standpoint in mind, Shipstead, Redick et al. (2012) examined two large datasets that included multiple measures of complex span and a color-change version of visual arrays. The analyses were performed using latent-variable structural equation modeling (Fig. 2). In this technique regressions are performed using factors that are formed by extracting variance that is common to several tasks,

Fig. 1 Examples of (a) visual arrays and (b) complex span tasks. In the visual arrays task the test-taker needs to decide whether the color of a box has changed, relative to its initial presentation. In the complex span task a series of items (e.g., letters) must be remembered, the presentation of which is interrupted by a simple processing task

thus reducing the influence of task-specific components. Fluid intelligence (novel problem solving ability) served as a means of validating these factors. That is, using latent regression, Shipstead, Redick et al. (2012) were able to demonstrate that visual arrays and complex span not only represent separate factors, but each factor is meaningful to complex cognition.

As can be seen in Fig. 2, the findings were straightforward and replicable. The factors formed by visual arrays (VA[1]) and complex span (CS) tasks were separable, but strongly correlated. This was consistent with our theory that these factors represent separable aspects of the same cognitive system (Shipstead, Redick et al., 2012).

Furthermore, each of these working memory factors was related to general fluid intelligence (Gf). The numbers on the lines between each working memory factor and Gf are standardized regression coefficients, and thus represent the correlation of one factor to Gf, when the other is held constant. As can be seen, in both datasets CS had a much stronger relationship to Gf than did VA. Shipstead, Redick et al. (2012) interpreted this as evidence that control processes that are apparent in complex span tasks play a larger role in relating working memory capacity to reasoning ability than does raw maintenance capacity, as reflected in visual arrays.[2]

Concerns with these findings

While the results of Shipstead, Redick et al. (2012) fit a reasonable narrative, there are several shortcomings related to the use of pre-existing data sets. First, these data sets only included one type of visual arrays task, but several types of complex span. In order to construct a proper structural equation model, Shipstead et al. were forced to divide the visual arrays task by set size (four, six, or eight items). This was done under the assumption that a common mechanism was driving performance in all of these sets, and it would be extracted in the structural equation model. Although this is technically correct, the tasks in the complex span factors were more varied in their demands (verbal vs. spatial memory; different types of processing tasks). The CS factor thus represented performance across a greater variety of contexts, and was therefore likely a more pure measure of the mechanisms that are critical to complex span performance, relative to VA. The visual arrays factor likely included more task-specific variance that was unrelated to fluid intelligence.

Second, the models of Shipstead, Redick et al. (2012) did not include any measures of attention control against which their arguments could be validated. This is a critical omission.
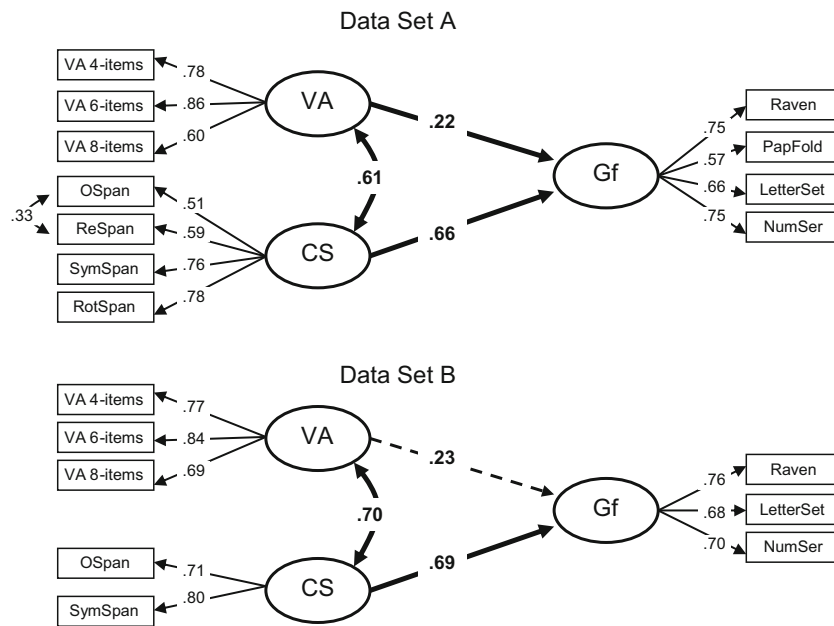
Although many researchers take as a pre-experimental given that visual arrays performance represents temporary storage in a memory buffer (Chuderski et al., 2012; Cowan et al., 2005; Shipstead, Redick et al., 2012; Unsworth et al., 2014), there are reasons to doubt this position. In particular a number of studies have revealed the presence of proactive interference in this task (Hartshorne, 2008; Shipstead & Engle, 2013; Souza & Oberauer, 2015; see also Endress & Potter, 2014). Protection from proactive interference is often considered to be a critical hallmark of temporary storage (Cowan, 2001). This suggests that visual array performance arises – at least in part – from effective control processes (e.g., attention control, controlled retrieval) that function to combat the effects of proactive interference.

Although contrary to traditional interpretation (cf. Cowan, 2001; Luck & Vogel, 1997), experimental results do indicate that visual arrays performance does involve a degree of controlled attention. For instance, Fukuda and Vogel (2009), 2011) have repeatedly demonstrated that individual differences in visual arrays performance predict a person's ability to recover from attention capture. That is, high performers on change detection tasks are also less likely to fall prey to environmental distraction. Moreover, they seem to be better able to filter out visual noise to retain access to only critical information (McNab & Klingberg, 2008; Vogel et al., 2005; but see Mall et al., 2014). This indicates that at least two attentional mechanisms may be present: (1) sustained attention control and (2) filtering of unimportant information. Cusack et al. (2009) suggest that such mechanisms allow people to restrict their maintenance of items in an array, and thus create strong encodings of a few items, rather than weak encodings of many items. Shipstead et al. (2014) demonstrated that, across a broad range of task demands, visual arrays performance has a strong correlation to traditional attention control tasks. A reanalysis of this data set along with a replication against another recently collected set will form the basis of the present study.

Testing the conclusions of Shipstead, Harrison et al. (2012) and Shipstead, Redick et al. (2012)

Test 1: the relationship of visual arrays to fluid intelligence

Since the publication of Shipstead, Redick et al. (2012) we have collected two large data sets as part of separate studies (Shipstead et al. 2014; Shipstead, Harrison, Trani et al., submitted for publication; Shipstead, Harrison, & Engle, submitted for publication). Each of these data sets improves on the originals by including multiple types of visual arrays tasks, as well as measures of attention control. Thus, we can create models such as those in Fig. 2 with greater fidelity, since task-specific variance will be eliminated in VA to a greater degree. Second, the inclusion of measures of attention control will allow us to not only examine the possibility that

---

[1] From this point forward, abbreviations will be used as reference to factors presented in our latent-variable models.

[2] Note that, although the path from VA to Gf was non-significant in data set B, the path sizes are comparable between data sets. This indicates that the non-significant path was likely a power issue, as data set B ($n = 170$) contained a much smaller sample than data set A ($n = 505$).

Fig. 2 Recreation of structural equation models reported by Shipstead, Redick et al. (2012). Circles represent latent factors: CS = complex span; Gf = fluid intelligence; VA = visual arrays. Boxes represent individual tasks: OSpan = operation span; NumSer = number series; PapFold = paper folding; Raven = Raven's Advanced Progressive Matrices; ReSpan = reading span; RotSpan = rotation span; SymSpan = symmetry span; numbers next to visual arrays tasks indicate the number of items shown in the display

control-processes are directly linked to both simple mainte-nance (i.e., no interruption) and complex maintenance (i.e., memory in spite of interruption) in working memory, but also examine the relative importance of attention control in each of these scenarios.

Test 2: the relationship of complex span and visual arrays to the control of attention

Figure 3 presents several ways of representing the relationship of working memory capacity to attention control (AC). Note that technically the arrows should be pointing from AC to WMva and WMcs, since our theory states that attention con-trol is a causal factor. However, treating attention control as the dependent variable allows us to better define the regres-sions that are being performed. Thus, we caution against drawing causal conclusions regarding these correlational models, they are designed with variance portioning in mind.

First, Fig. 3a represents our original perspective (Shipstead, Redick et al., 2012) and a perspective that is con-sistent with the assumption that the focus of attention is pri-marily a storage system, or memory buffer. In this model, WMva and WMcs are correlated, which we interpret as a product of these factors representing two qualitatively inter-active components of the same system (scope and control of attention). This model assumes that WMcs is exclusively re-lated to AC. WMva only relates to AC to the extent that it shares variance with WMcs. It does not add to the prediction of attention control. This outcome was predicted by Shipstead, Redick et al. (2012).

Second, Fig. 3b represents the opposite solution. It pro-poses that the size of focal attention is strongly related to attention control. In this model, WMcs has no direct relation-ship to AC. Instead, WMcs is only related to AC to the extent that it reflects the focus of attention. From the perspective of Shipstead, Redick et al. (2012) this outcome is highly
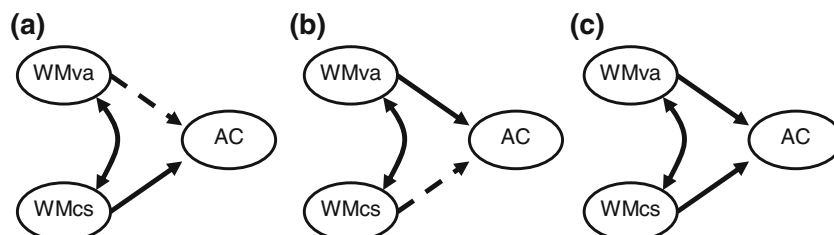
Fig. 3 Predicted result, for cases in which (**a**) the relationship between working memory as measured by visual arrays (WMva) and attention control (AC) is mediated by working memory capacity as measured by complex span (WMcs), (**b**) the relationship between WMcs and AC is mediated by WMva, or (**c**) both measures of working memory capacity have relationships to attention control above-and-beyond one another

unlikely. However, given the possibility that visual arrays requires a strong component of attention control (Fukuda & Vogel, 2009, 2011; Shipstead et al., 2014), such a model is possible. It would, in principle, reinforce recent claims that attention control largely determined by maintenance capacity (Chuderski et al., 2012). We will explore the difficulties associated with interpreting such a relationship, following the main analyses.

Third, Fig. 3c represents a solution in which both WMcs and WMva have relationships to AC above and beyond one another. That is to say, both measures of working memory capacity uniquely represent variance associated with attention control. Thus, even when variance associated with complex span performance is accounted, the size of focal attention retains a relationship to attention control. This outcome would force a reinterpretation of the work of Shipstead, Redick et al. (2012) which treated WMcs as an indicator of working memory capacity processing and WMva as an indicator of working memory capacity storage. However, despite its incongruence to our theory, we did not judge this outcome to be unlikely, given recent studies by Fukuda and Vogel (2009, 2011) and Cusack et al. (2009). Figure 3c would be interpreted as indicating that processes such as attention control are an important determinant of working memory capacity, regardless of the type of task that is used, and perhaps an important determinant of the effective size of attentional capacity.

Method

Data sets

Analyses were performed using two previously collected data sets. Data Set 1 includes data originally reported in Shipstead et al. (2014). This sample included 215 people from the Atlanta, GA, USA community (48 % female; 60 % college students) between the ages of 18 and 30 years. Data Set 2 includes data that was originally reported in Shipstead, Harrison, Trani et al. (submitted for publication) and Shipstead, Harrison, & Engle (submitted for publication). This sample included 573 people from the Atlanta, GA and Columbus, IN communities (47 % female; 62 % college students) between the ages of 18 and 35 years. Both experiments were multisession and participants were run in groups of 1–5. Each session lasted 2 h. The order in which tasks were run can be found in Appendix A (Table 5).

Tasks

Working memory capacity (visual arrays)

All visual arrays tasks presented test-takers with an array of items on a computer screen that briefly disappeared, then returned (see Fig. 1a). The test-taker's job was to indicate whether an aspect of the display had changed relative to the first presentation. Changes occurred on half of all trials. k served as the dependent variable for each task. Per the recommendations of Rouder et al. (2011), k was calculated using one of two formulas. On tasks in which test-takers responded as to whether a probed item had changed, the formula was $k = N * (\text{hits} + \text{correct rejections} -1)$. On tasks in which test-takers responded as to whether any item on the screen may have changed, the formula was $k = N * (\text{hits} - \text{false alarms} / (1 - \text{false alarms}))$. For visual arrays tasks in which all initially displayed items were relevant, N was the number of items in the array. For visual arrays tasks that required attentional filtering of distractor items, N was defined as the number of to-be-attended items on the screen.

At a distance of 45 cm from the monitor, the items were presented within a 19.1° × 14.3° field. All items randomly presented with the constraint that they were separated by at least 2° and were 2° from central fixation. This is similar to tasks used by Fukuda et al. (2010) and Shipstead and Engle (2013). For color change tasks, white, black, red, yellow, green, blue, and purple were used, and the colors could repeat on a screen. Squares subtended 0.09° of visual angle along any side. For orientation change tasks, red and blue items were shown in equal numbers. Rectangles subtended 0.32° of visual angle, along the wide edge.

VAcolor (Data set 1; data set 2) Arrays were composed of 4, 6, or 8 colored boxes. Colors could repeat within a given array and included white, black, red, yellow, green blue, and purple. The arrays were presented for 250 ms, followed by a 900-ms retention interval. Memory was probed by circling one item in the test display. Test-takers needed to decide whether this item had changed color. For each set size, 28 trials were run.

VAorient (Data set 1; data set 2) Arrays were composed of five or seven colored bars (blue or red) that were horizontal, vertical, or 45° to the left or right. Timing characteristics matched those of VAcolor. Test-takers needed to decide if any item in the test display had changed orientation. For each set size, 40 trials were run.

VAcolorS (Data set 1) Each trial on this task began with an arrow that pointed left or right (100 ms), and indicated which side of the screen should be remembered. After 100 ms arrays of four, six, or eight items appeared on the left and right sides of the screen. Display time of these arrays was reduced to 100 ms in order to reduce concern that the average test-taker could reorient his/her eyes with longer presentations. After a 900-ms retention interval the boxes returned on the side to which the arrow had pointed. Test-takers needed to decide if any item had changed color. For each set size, 28 trials were run.

VAorientS (Data set 1; data set 2) Each trial began with the word "red" or "blue" (200 ms) indicating that test-takers should only attend to red or blue items in the subsequent display. After a 100-ms display, an array of either 10 or 14 bars at different orientations (see VAorient) were displayed for 250 ms. Half of the bars matched the color that had been given at the beginning of the trial. After a 900-ms retention interval, the to-be-attended bars returned. Memory was probed by superimposing a white dot on one of the bars. Test-takers needed to decide whether this item had changed orientation. Forty trials of each set size were run.

Working memory capacity (complex span)

All complex span tasks were computer based tasks (Unsworth et al., 2005; Unsworth, Redick et al., 2009) in which test-takers remember a series of sequentially presented items (see Fig. 1b). This sequence was interrupted by a simple processing task that needed to be performed before the next item was presented. Lists of items varied in length and each list length was presented three times. In Data Set 1, the list lengths were randomized. In Data Set 2, test-takers performed three blocks of trials with each list length presented once per block. List length presentation was random within each block (see Foster et al., 2015 for more information on these tasks). The dependent variable was the number of items recalled in their correct serial position.

OSpan (Data set 1; data set 2) The to-be-remembered items were letters. The interpolated processing task was a simple math problem. List lengths varied from 3–5 items.

SymSpan (Data set 1; data set 2) The to-be-remembered items were spatial locations in a 4 × 4 grid. The processing task was a symmetry judgment regarding a black and white figure laid out in a 8 × 8 grid. List lengths varied from 2–5 items.

RotSpan (Data set 2) The to-be-remembered items were the directions of long and short arrows radiating from a central location. The processing task required test-takers to decide if a rotated letter was normally-oriented or mirror-reversed. List lengths varied from 2–5 items.

Attention control

AntiSac (Data set 1) The antisaccade task (see Hutchison, 2007) required test-takers to divert their gaze from a peripheral flash that occurred on one side of a computer monitor and report a letter (O or Q) that was briefly presented on the opposite side of the computer monitor. This sequence of events began with a fixation-point (+) that was displayed for either 1000 or 2000 ms. The to-be-reported letter was presented for

100 ms and then masked by "##". 5000 ms were given to report the displayed letter via key press. The dependent variable was accuracy over 48 trials.

BeepSac (Data set 2) The beep-antisaccade (see Shipstead, Harrison, Trani et al., submitted for publication; Shipstead, Harrison, & Engle, submitted for publication) was the same as the AntiSac with the exception that a short beep played 300 ms before the peripheral flash.

Antisac2 (Data set 2) Each trial of the antisac2 task (see Kane et al., 2001) began with a "***" fixation-point on a computer monitor that varied in duration (200–1800 ms). This was followed by a "=" that flashed twice (300 ms) on either the right or left hand side of the screen. Next, a letter was presented on the opposite side of the screen for 100 ms then masked by the number "8". Test-takers were allowed 10,000 ms to report whether the masked letter had been "B", "P", or "R". The dependent variable was accuracy over 60 trials.

Stroop task (Data set 1; data set 2) The Stroop (1935) task required test-takers to report the hue in which a color-word was presented via keypress. Blue, green, and red were used and corresponding color-stickers were placed on the number pad of the keyboard. On 66 % of trials the word and hue were congruent (half were used in data analyses and half were fillers). The remaining 33 % were incongruent. For Data Set 1 there were a total of 486 trials. For Data Set 2 there were a total of 162 trials. The dependent variable was response time differences between congruent and incongruent trials.[3]

Flanker (Data set 1; data set 2) The arrow flanker task required test-takers to report whether a central arrow was pointing left or right. Flanking arrows could be congruent (→ → → → →), incongruent (← ← → ← ←), or neutral (— — → — —). Each combination was equally probable, and 216 trials were run. The dependent variable was response time differences between incongruent and neutral trials.

Fluid intelligence

Raven (Data set 1; data set 2) On each trial of Raven's Advanced Progressive Matrices (Raven, 1990; odd problems), eight abstract figures were displayed in a 3 × 3 matrix. The final position was blank. Test-takers select one of several options that belonged in the missing space. The dependent variable was the number of problems solved in 10 min (18 total).

---

[3] When reported in Shipstead, Harrison, Trani et al. (submitted for publication; Shipstead, Harrison, & Engle, submitted for publication) the Stroop and Flanker tasks' dependent variables used a method of Hughes et al. (2013) that estimated a speed accuracy tradeoff. In order to increase consistency across data sets, we used the classic dependent variable.

LetterSet (Data set 1; data set 2) On each trial of letter sets (Ekstrom et al., 1976), five four-letter strings were presented. Four of the sets followed the same rule. The test-taker needed to discover the rule and select the string that did not follow it. The dependent variable was the number of correct responses given in seven minutes (30 total).

NumSer (Data set 1; data set 2) In the number series task (Thurstone, 1938) a series of numbers was presented on a computer screen. These numbers were joined by a rule that the test-taker needed to discover. The response involved deciding what the next number in the series would be. The dependent variable was the number of correct responses provided in five minutes (15 problems).

Fit statistics

Multiple fit indicators are reported for each model. χ^2/df (chi-square divided by degrees of freedom) is a "badness-of-fit" statistic. Important to this statistic, χ^2 is sensitive to sample size, while df is determined by the number of parameters in a model. Thus there can be no hard-and-fast rule for interpreting this statistic (Kline, 1998). Because of sample size differences we considered values up to 2 to be adequate for Data Set 1 and values up to 3 to be adequate for Data Set 2. Root mean square error of approximation (RMSEA) provides an estimate of model fit to the population. Standardized root mean square residual (SRMR) provides average deviation of the covariance matrix produced by the model relative to the observed matrix. For both of these indices, values below .05 are preferred, but up to .08 are acceptable (Browne & Cudeck, 1993; Kline, 1998). Non-normed fit index (NNFI) and comparative fit index (CFI) compare the hypothesized model to a model in which the observed variables are assumed to be uncorrelated. In both cases, values above .95 are considered a good fit (Hu & Bentler, 1999). AIC (Akaike, 1987) is an indicator of model parsimony. It takes both goodness-of-fit and number of estimated parameters into account. When a path is added and AIC drops, it can be stated that increased explanatory power offsets the loss of parsimony.

Results and discussion

Descriptive statistics and correlation matrices can be found in Tables 1 and 2. Note that visual arrays task performance was not reported in Shipstead, Harrison, Trani et al. (submitted for publication) and Shipstead, Harrison, & Engle (submitted for publicationb). A correlation matrix that integrates the present data with the data of Shipstead et al. will be made available at http://psychology.gatech.edu/renglelab/omnibus_data

There are two points to be made regarding the visual arrays data on Table 1. First, the mean k values (the estimate of attention capacity drawn from visual arrays performance) for VA3 and VA4 are noticeably lower than the traditional score of 3–4. We interpret this as evidence of test-takers being imperfect in their ability to filter trial-irrelevant information. In essence, these people maintain access to information that is inconsequential to the eventual test. The effective result is a lower k value.

Second, on several of the visual arrays tasks, test-takers had extremely low k values (less than −1). One interpretation is that these test-takers confused their keys and thus were responding in reverse. This would make them candidates for removal from the data set. We examined the data for test-takers who had k values that were consistently lower than −1. None fit this criterion. In keeping with previous studies (Shipstead et al., 2014; Shipstead, Harrison et al., 2012; Shipstead, Redick et al., 2012), we thus interpret negative k scores as reflecting an inability to accurately maintain the array when it is not present. These scores differ from zero because the test-taker is responding based on little-to-no signal, which leads to guessing, which does not always balance out within a session. Scores that differ greatly from zero represent bad luck of the type that leads to regression to the mean.

In subsequent models Data Set 1 will include a correlated error term between the standard color change visual arrays task (VAcolor) and the selective orientation change task (VAorientS). This was expected based upon a previous analysis of the data (Shipstead et al., 2014) and a few points are important. First, there is a weaker correlation error between the error terms of the other two visual arrays tasks. However, this is only apparent if the included correlation is removed. Second, in Data Set 2 there is a non-significant negative error between the same tasks. This suggests that including both non-selective and selective visual arrays introduces an influence that works against the general visual arrays factor (likely visual filtering, which is unrelated to visual arrays or other measures of working memory capacity; Shipstead, Harrison et al., 2012; Shipstead, Redick et al. 2012; see also Mall et al., 2014). The negative error term for Data Set 1 was thus retained to remove this influence.

Relationship of working memory capacity to fluid intelligence

The first analysis replicated the models of Shipstead, Harrison et al. (2012) and Shipstead, Redick et al. (2012). In both cases the fit was good (Table 3: WMCtoGF). In examining the models (Fig. 4), several points are noteworthy. First, there is a reasonably strong correlation between WMcs and WMva. This is less so for Fig. 4a, relative to other models (Fig. 4b; Fig. 1). Nonetheless, in keeping with our contention that visual arrays and complex span measure related components of

Table 1 Descriptive statistics

Task	M	SD	Range	Skew	Kurtosis	I.C.
Data set 1						
1. OSpan	56.11	13.64	9.00–75.00	−.94	.66	.84[a]
2. SymSpan	26.46	8.74	3.00–42.00	−.50	−.30	.84[a]
3. VA1color	3.52	1.18	−.65–5.71	−1.17	1.86	.78[a]
4. VAorient	3.04	1.34	−1.66–5.45	−.73	.55	.74[a]
5. VAcolorS	2.01	1.44	−3.31–4.91	−.63	.76	.54[a]
6. VAorientS	1.66	1.23	−.80–4.88	.15	−.39	.70[a]
7. AntiSac	.74	.15	.21–1.00	−.67	−.16	.85[a]
8. Flanker	96.88	49.23	12.73–273.52	1.23	1.74	.81[b]
9. Stroop	138.96	85.37	−39.66–453.54	.90	.86	.92[b]
10. Raven	8.92	3.77	1.00–17.00	−.24	−.82	.80[a]
11. LetterSet	15.12	4.54	3.00–25.00	−.38	−.22	.82[a]
12. NumSer	8.73	3.08	1.00–15.00	−.37	−.22	.76[a]
Data set 2						
1. OSpan	54.17	15.51	.00–75.00	−.88	.19	.86[a]
2. SymSpan	26.66	9.05	3.00–42.00	−.43	−.49	.84[a]
3. RotSpan	24.62	9.77	.00–42.00	−.44	−.57	.87[a]
4. VAcolor	3.37	1.46	−1.24–5.76	−.91	.36	.84[a]
5. VAorient	2.69	1.60	−3.04–5.77	−.78	.84	.73[a]
6. VAorientS	1.58	1.30	−1.68–5.00	.12	−.48	.79[a]
7. BeepSac	.79	.16	.29–1.00	−.70	−.49	.91[b]
8. AntiSac2	.55	.16	.20–.98	.12	−.80	.81[b]
9. Stroop	132.26	98.89	−152.63–479.20	.71	1.03	.60[b]
10. ArrowFlanker	99.69	45.49	−49.97–260.85	.75	1.38	.66[b]
11. Raven	8.69	3.91	.00–18.00	−.07	−.88	.82[a]
12. LetterSet	15.38	5.29	1.00–29.00	−.03	−.62	.84[a]
13. NumSeries	8.56	3.58	.00–15.00	−.22	−.86	.83[a]

[a] Internal consistency calculated using Chronbach's α

[b] Internal consistency calculated using split-half

working memory, WMva and WMcs do have a strong relationship, on the whole.

Second, the relationships of WMva and WMcs to Gf are more balanced than in the models of Shipstead, Harrison et al. (2012) and Shipstead, Redick et al. (2012) (Fig. 1). Indeed, Fig. 4 indicates that, holding one factor constant, WMva and WMcs each have an equally strong relationship to Gf. In other words, contrary to previous results, simple maintenance (as reflected by visual arrays) is equally critical to fluid reasoning ability as is complex maintenance that occurs in the face of interruption (as reflected by complex span). We directly verified these statements by forcing the paths from WMva and WMcs to Gf to be equal in both models. As can be seen in Table 3 (WM to Gf – Equal), this did not disrupt the fit of either model.

The disparity between the present results and those of Shipstead, Redick et al. (2012) is easily reconciled. The visual arrays factor of Shipstead, Redick et al. (2012) only included one type of visual arrays task, while the present models included 3–4 distinct tasks. Thus, in the present models, the tasks that formed WMva were more varied in their demands, and thus task-specific variance was better excluded from WMva. In other words, our current measure of the scope of attention was a better reflection of the processes that are critical to maintenance across a variety of visual arrays tasks, as compared to what was reflected by a single task.

As such, the proper conclusion to draw is that the present models provide a more valid measure of maintenance in focal attention. This observation necessitates an update of our thinking about our previous findings. As it turns out, the cognitive mechanisms that are reflected in visual arrays performance are equally important to explaining the relationship between working memory capacity and fluid intelligence as is are those reflected in complex span tasks. However, as is made clear by subsequent models, the maintenance processes reflected in visual arrays tasks represents more than transient storage.

Table 2   Correlations among all tasks

Task	1	2	3	4	5	6	7	8	9	10	11	12	13
Data set 1													
1. OSpan	–												
2. SymSpan	.52	–											
3. VAcolor	.30	.39	–										
4. VAorient	.27	.38	.59	–									
5. VAcolorS	.21	.32	.50	.42	–								
6. VAorientS	.23	.36	.44	.59	.54	–							
7. AntiSacc	.23	.40	.41	.42	.44	.45	–						
8. Flanker	−.18	−.23	−.25	−.21	−.25	−.22	−.28	–					
9. Stroop	−.17	−.24	−.12	−.09	−.14	−.22	−.13	.23	–				
10. Raven	.34	.49	.45	.41	.39	.43	.44	−.23	−.07	–			
11. LetterSet	.29	.41	.36	.34	.30	.37	.37	−.09	−.10	.54	–		
12. NumSer	.30	.41	.38	.37	.31	.36	.36	−.15	−.08	.58	.54	–	
Data set 2													
1. OSpan	–												
2. SymSpan	.54	–											
3. RotSpan	.53	.68	–										
4. VA Color	.43	.46	.48	–									
5. VA Orient	.35	.42	.41	.59	–								
6. VA Select	.44	.50	.51	.61	.58	–							
7. BeepSaccade	.39	.48	.49	.47	.47	.53	–						
8. AntiSaccade	.38	.44	.44	.43	.39	.50	.61	–					
9. Stroop	−.29	−.19	−.18	−.22	−.17	−.21	−.21	−.22	–				
10. ArrowFlanker	−.09	−.14	−.18	−.17	−.21	−.18	−.18	−.21	.11	–			
11. Raven	.51	.36	.59	.56	.53	.58	.54	.45	−.23	−.23	–		
12. LetterSet	.46	.27	.55	.54	.41	.54	.51	.50	−.22	−.22	.61	–	
13. NumSeries	.55	.32	.56	.58	.45	.53	.49	.50	−.29	−.24	.65	.68	–

Complex span, visual arrays, and attention control

As we indicated in the introduction (see also, Shipstead, Harrison et al., 2012; Shipstead, Redick et al., 2012) it may be misleading to refer to WMva and WMcs as the "scope" and "control" of attention. In particular, this gives the impression that focal attention is a memory buffer in which information is temporarily stored. In fact, the scope of attention may arise from complex processes approximating attention control. Working memory maintenance may reflect the efficacy with which control processes function, even when interruption is minimal (Cusack et al., 2009; Fukuda & Vogel, 2009, 2011; Vogel, McCollough, & Machizawa; Shipstead & Engle, 2013).

The next set of analyses tested this possibility by relating WMva and WMcs to attention control (AC) using the model from Fig. 3c. Fit statistics are located on Table 4.

Table 3   Fit statistics for models relating working memory capacity to fluid intelligence

Model	χ2	df	χ2/df	RMSEA	SRMR	NNFI	CFI	AIC
WM to Gf Data set 1	8.51	23	.37	.00	.02	1.02	1.00	52.51
WM to Gf Data set 2	70.81	24	2.95	.06	.03	.99	.99	112.81
WM to Gf Data set 1 - Equal	8.63	24	.36	.00	.02	.99	1.00	50.63
WM to Gf Data set 2 - Equal	70.82	25	2.83	.06	.03	.99	.99	110.82

RMSEA root mean square error of approximation, *SRMR* standardized root mean square residual, *NNFI* non-normed fit index, *CFI* comparative fit index, *AIC* indicator of model parsimony
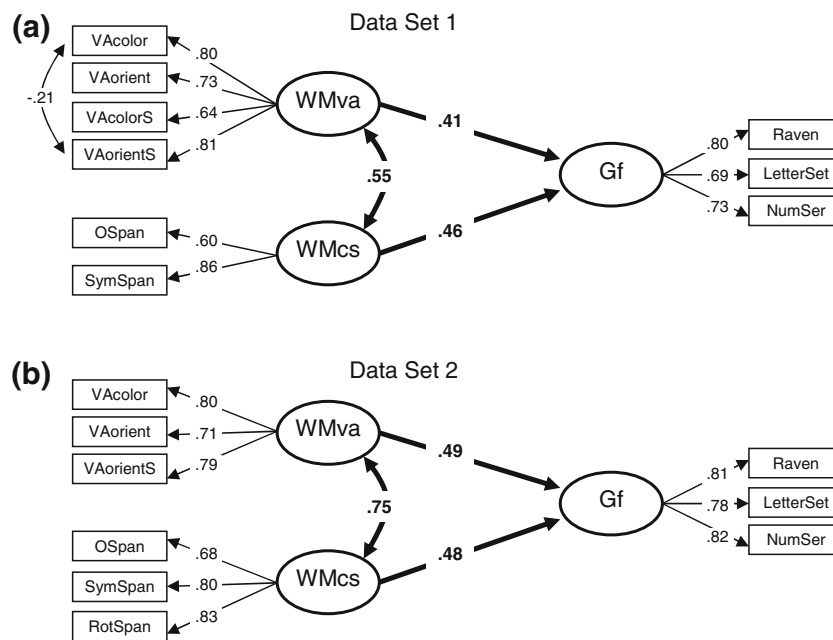
Fig. 4 Structural equation models relating working memory capacity, as reflected in visual arrays performance (WMva) and complex span (WMcs) to fluid intelligence (Gf). OSpan = operation span; LetterSet = letter sets; NumSer = number series; Raven = Raven's Advanced Progressive Matrices; RotSpan = rotation span; SymSpan = symmetry span; VAcolor = visual arrays, color change; VAorient = visual arrays orientation change; VAcolorS and VAorientS are selective filtering versions

Figure 5 reveals highly consistent results between the models. In both data sets WMva and WMcs have significant relationships to AC, above-and-beyond each other and the regression paths are strikingly similar between models. In fact, in both cases WMva actually has a numerically larger relationship to AC than does WMcs. This outcome clearly falsifies the assertion made by Shipstead, Harrison et al. (2012) and Shipstead, Redick et al. (2012) that WMva represents the scope of attention, while WMcs represents the control of attention. Visual arrays tasks and complex span tasks each predict unique variance in attention control (see also Cowan et al., 2006).

Unexpectedly, WMva had a numerically larger path to AC than did WMcs. We next tested whether visual arrays should

be considered the stronger predictor of attention control by constraining the two path models such that the models were required to arrive at a solution in which WMva and WMcs had equivalent paths to AC. The idea is that, if WMva is the stronger predictor, the models labeled "Equal Paths" (Table 4) would have a poor fit, relative to the models labeled "Both to AC".

For Data Set 1, this constraint resulted in paths of .49 from WMcs and WMva to AC. For Data Set 2, this resulted in paths of .44. In neither case was the model labeled "Equal Paths" noticeably disruptive to fit, relative to "WM to AC" (Table 4). In Data Set 2, the constraint leads to a somewhat inflated χ^2 and a slight increase in AIC. This indicates that it is not necessarily

Table 4 Fit statistics for models relating working memory capacity to attention control

Model	χ2	df	χ2/df	RMSEA	SRMR	NNFI	CFI	AIC
WM to AC Data set 1	24.60	23	1.07	.02	.04	1.00	1.00	68.60
WM to AC Data set 2	50.88	32	1.59	.03	.03	.99	1.00	96.88
Equal Paths Data set 1	25.86	24	1.08	.02	.04	.99	1.00	67.86
Equal Paths Data set 2	53.87	33	1.63	.03	.03	.99	.99	97.87
Miscellaneous models								
Appendix B Model Data set 1	12.14	11	1.13	.02	.04	1.00	1.00	46.41
Appendix B Model Data set 2	42.82	24	1.78	.04	.03	.99	.99	84.82

RMSEA root mean square error of approximation, *SRMR* standardized root mean square residual, *NNFI* non-normed fit index, *CFI* comparative fit index, *AIC* indicator of model parsimony
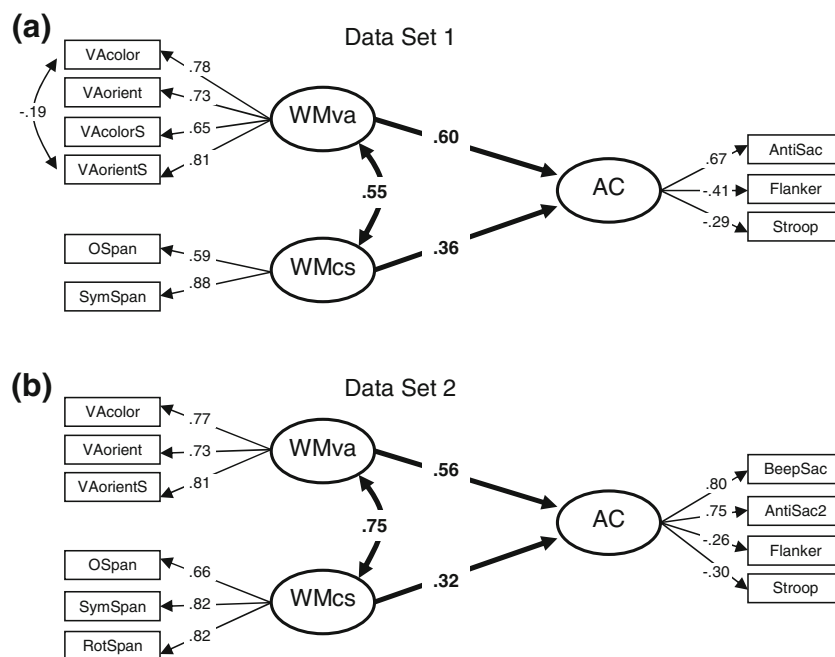
Fig. 5 Structural equation models relating working memory capacity, as reflected in visual arrays performance (WMva) and complex span (WMcs) to attention control (AC). AntiSac/AntiSac2 = antisaccade; BeepSac = antisaccade with warning beep; Flanker = flanker task; OSpan = operation span; RotSpan = rotation span; SymSpan = symmetry span; VAcolor = visual arrays, color change; VAorient = visual arrays orientation change; VAcolorS and VAorientS are selective filtering versions

appropriate to assume the paths are exactly equal. However, on the whole the model fits are not substantially disrupted by assuming that WMva and WMcs have equivalent relationships to AC. We therefore caution against concluding that visual arrays performance is a stronger indicator of attention control than is complex span performance. Instead, we simply conclude that each of these tests of working memory capacity provides a strong reflection of a person's attention control abilities, and these tests reflect both common and unique aspects of attention control.

General discussion

The present study updates previous work on the distinction between working memory as measured by simple change detection tasks (visual arrays) and complex span tasks, which force people to remember information in the face of distraction (Cowan et al., 2005; Shipstead, Harrison et al., 2012; Shipstead, Redick et al., 2012). Although these tasks converge on separate factors, the factors are strongly related. In keeping with the demands that these tasks make on the working memory system, we began by respectively interpreting them as reflecting (1) moment-to-moment maintenance capacity, and (2) maintenance that occurs during periods of interruption (or as the result of controlled retrieval).

In terms of the relationship between working memory capacity and fluid intelligence, the present results clearly contradicted the conclusions of Shipstead, Harrison et al. (2012) and Shipstead, Redick et al. (2012) (Fig. 1). Complex span performance was not found to have the stronger relationship to fluid intelligence. Instead, the relationship, at our most conservative estimate, was balanced. The latent factors underlying complex span and visual arrays performance provided roughly equivalent prediction of fluid intelligence. Rather than interpreting this as between-study inconsistency (relative to Shipstead, Harrison et al., 2012; Shipstead, Redick et al., 2012), we instead note the importance of including diverse tasks in any measure of ability. Failure to do so will create a factor with a high amount of task-specific variance, which may have an attenuating effect on its relationship to other constructs.

The more surprising finding was the strong relationship of visual arrays to attention control, even when variance associated with complex span performance was controlled. Previous studies have concluded that visual arrays performance is predictive of attention control (Fukuda & Vogel, 2009, 2011; Shipstead et al., 2014); however, as pointed out in our hypothetical models (see Fig. 3a), this might have been explained by the relationship of WMva to WMcs. If both of these factors represent components of the same cognitive system, then the correlation between visual arrays
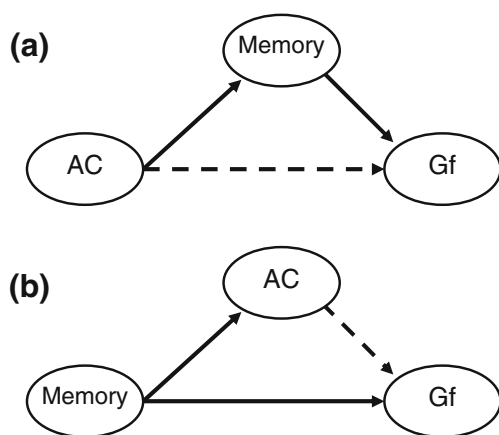
Fig. 6 Hypothetical models in which (**a**) memory mediates the relationship between attention control and fluid intelligence, and (**b**) memory is a common cause of attention control and intelligence that fully accounts for the relationship between these variables. Although these models express qualitatively different relationships, they are actually mathematically equivalent

performance and attention control could be mediated by processes that are critical to complex span. In such a case, holding WMcs constant would eliminate the relationship between WMva and AC. This expectation did not hold. In fact, visual arrays tasks, which require only simple maintenance of information, may have a stronger relationship to attention control than does complex span performance.

Visual arrays, focal attention, and attention control

An obvious counter argument to the claim that visual arrays performance consistently reflects a person's attention control is that some of the visual arrays tasks included an attention filtering requirement, which may have introduced attention control demands that are not otherwise present. There are two important points to be made. First, the standard visual arrays tasks had factor loadings that were roughly equivalent to the selective versions. Thus, they were contributing equally to the overall factor.

Second and related to this point, if the tasks are contributing equally, the correlation of the visual arrays factor to attention control should be robust to the removal of selective tasks. We put the idea to the test by creating models that only included the two simple change-detection visual arrays tasks. These models are located in Appendix B and fit statistics are in the portion of Table 4 labeled "Appendix B Model." As can be seen, these analyses conformed to our main findings. The strong correlations to attention control remained.

This relationship is interesting, in and of itself, because it is counterintuitive. Why should a task that has been long labeled as measuring the size of momentary

visuo-spatial storage be so strongly related to the ability to control attention? Of course, due to the limitations of strict correlational approaches, we are limited as to the statements we can firmly support within the present data sets. There are, however, two ways of interpreting this relationship.

Figure 6 presents two ways that the relationship between memory, attention control, and fluid intelligence can be represented. Examination of the source articles from which the present data were drawn (Shipstead, Harrison, Trani et al. submitted for publication; Shipstead, Harrison, & Engle, submitted for publicationb; Shipstead et al., 2014) as well as other studies (Chuderski et al., 2012; Unsworth & Spillers, 2010; but see Unsworth, Spillers et al., 2009) reveals that a point expressed in both of these figures is accurate: The correlation between attention control and fluid intelligence is largely explained by memory-related factors.

Figure 6a presents the data from our perspective. Attention control is responsible for ensuring that a person is oriented toward maintaining relevant information. In turn, this facilitates reasoning. In this model, "memory" does not represent the capacity of a buffer, so much as it represents performance on short-term memory tasks that is facilitated by the ability to stabilize attention on relevant information.

Figure 6b presents a perspective similar to one proposed by Chuderski et al. (2012). In this model, temporary storage is seen as a common cause of attention control and fluid intelligence. Once storage-related processes are accounted, the correlation between attention control and fluid intelligence disappears.

These are two qualitatively different conclusions. However, these models are actually mathematically equivalent. They will produce the same path coefficients and the same fit statistics.

Ultimately, support for our position or the position expressed Chuderski et al. (2012) suffer from the same limitations. While correlational work can uncover unexpected relationships, it rarely demonstrates causality. It is thus important that we support our interpretation of the above models with experimental observations that indicate visual arrays performance (or change detection in general) is not a pure measure of temporary storage capacity.

First, one of the assumed hallmarks of short-term storage is protection from proactive interference (Cowan, 2001; Cowan et al., 2005). Thus, one would expect that the k-values drawn from visual arrays performance would be relatively invulnerable to proactive interference. However, using a temporal discriminability manipulation, Shipstead and Engle (2013) repeatedly demonstrated that effects of proactive interference are observed in this task, even with sub-capacity arrays sizes (2–3 items). When

two trials were run with relatively little time between them, k decreased (i.e., time-based build-up of proactive interference). When two trials were separated by relatively long delays, k increased (i.e., time-based release from proactive interference). In short, a person's ability to discriminate between information that was presented on the current trial and information that was presented on trial n-1 is a determinant of performance on this task. This would not be the case if visual arrays performance strictly reflected storage in a protective memory buffer (for further examples see Hartshorne, 2008; Makovski and Jiang, 2008; Souza & Oberauer, 2015).

Second, turning specifically to attention control, Fukuda and Vogel (2009, 2011) have repeatedly demonstrated that individual differences in visual arrays performance predict a person's ability to recover from attention capture. That is, people with higher k scores recover more quickly from having their attention drawn by irrelevant visual information. Our own interpretation of this finding is that it reveals an important mechanism underlying an individual test-taker's eventual k score. People with strong attention control stay focused on the task, rather than being drawn into random events in the environment or by their own thoughts. The lower the probability of succumbing to distraction at encoding or maintenance, the higher the probability of being able to accurately recognize that an object in the display has indeed changed.

Third, and at the heart of the distinction between the models in Fig. 6, attention control tasks simply do not place a heavy burden on maintenance (Roberts et al., 1994). They only require a test-taker to remember a simple to-be-performed behavior. People who perform poorly on attention control tasks do so because they (1) are deficient in the ability to simply maintain access to an instruction, (2) are deficient in the ability to retrieve the instruction when access is lost, and (3) also have difficulty dealing with response competition (see Kane & Engle, 2003).

Finally, many studies have shown that working memory is critical to setting and maintaining processing priorities for attention (Konstantinou & Lavie, 2013; Lavie, 2005; Lavie et al., 2004) and attention is often biased toward processing information that resembles maintained representation (Soto et al., 2005, 2008). The contents of working memory are the priorities of attention, and this relationship is more obvious when memory load is low (e.g., Soto et al., 2005) than when it is high (e.g., Kim, Kim, & Chun, 2005; Lavie et al., 2004). Working memory capacity is not simply a matter of how much information a person can maintain, but the ability to put that information to use. Extending the present findings, attention control provides a mechanism through which

maintained information can be protected for various forms of distraction that might overwrite processing goals.

From our perspective, k is less of a measure of maximum maintenance capacity, as it the ability to keep attention focused on the task at hand. Consistent focus leads to higher accuracy and higher k values are a byproduct. Furthermore, recent evidence suggests this is not limited to encoding and maintenance, but also extends to the testing phase.

Specifically, relative to change detection, visual arrays accuracy is diminished when the test requires 2-alternative-forced-choice recognition (i.e., which of two colors was in this location; Makovski et al., 2010). This indicates that active maintenance is less stable than suggested by the classic interpretation of k: The maintained representation of the target array is not robust to interference from the probe, and new input at test leads to greater interference. This implies yet another point in processing where attention control may be important. If attending to the test array decreases the accessibility of the maintained target array, then control processes provide a mechanism for balancing between maintaining the target and attending to the probe. Although a definitive test needs to be run, we predict that people with higher attention control will be less likely to confound the target and probe arrays, thus showing less sensitivity to testing method.

Multiple mechanisms of visual arrays performance

We began this article by describing the classic- and still fairly standard -explanation of visual arrays performance in which k represents the number of discrete items a person can simultaneously maintain in working memory. For reasons outlined in the preceding section, we favor a perspective in which attentional control is the more fundamental mechanism. At the same time, attention control does not provide a full account of visual arrays performance. Analysis of the models in Fig. 5 reveals that these constructs share about 50 % of their variance (square the regression between WMva and AC, then add to that the result of multiplying correlation between WMva and WMcs by the regression between WMcs and AC). Although attention control is critical to visual arrays performance, there is substantial room for other mechanisms to account for individual differences.

Thus, the present study is a challenge to single-mechanism accounts of working memory, but does not challenge debate regarding (for instance) whether certain aspects of maintenance are determined by a fixed capacity (Awh et al., 2007) or flexibly distributed resource (Ma, Husain, & Bays, 2014), or whether or not the focus of attention is free of proactive interference (Cowan

et al., 2005 vs. Shipstead & Engle, 2013; Souza & Oberauer, 2015). Instead, such debates may well represent aspects of visual arrays performance that function along side attention control. Similar to a recent argument that random degradation of maintained items accounts for a least a portion performance (Fougnie, Suchow, & Alverez, 2012), we believe that visual arrays performance (and working memory in general; Shipstead et al., 2014; Unsworth et al., 2014) can no longer be parsimoniously accounted through one mechanism.

The focus of attention

Shipstead, Redick et al. (2012) hypothesized that fundamentally different processes underlie visual arrays and complex span performance. Since we now argue that visual arrays performance strongly represents the control of attention, it is perhaps important to address our perspective on the size of focal attention.

As with visual arrays performance, we believe that complex span performance relates to focal attention because it represents a person's ability to stabilize focal attention around critical information, and to accurately recall information that has been displaced (Shipstead et al., 2014; Unsworth & Engle, 2006b; Unsworth et al., 2014). A person who performs these actions effectively will, for all intents and purposes, have a relatively large focus of attention. Indeed, Shipstead et al. (2014) found that controlling for individual differences in attention control caused the correlation between working memory capacity (as measured by either complex span or visual arrays) and verbal maintenance capacity to disappear.

At the same time, we do not intend to be dismissive of the concept of individual differences in the size of focal attention. Shipstead et al. (2014) also concluded that focal attention mediates any relationship between attention control and fluid intelligence. They thus concluded that focal attention represents more than working memory-related attention control and the additional factors underlying individual differences in the size of focal attention cannot be discerned by studying working memory capacity.

Interestingly, if focal attention does mediate the relationship between working memory capacity and fluid intelligence (Shipstead et al., 2014), it opens the possibility that certain cognitive mechanisms that are critical to reasoning (but are unrelated to working memory) may contribute to the explaining the size of a person's attentional focus. We suggest that a complete account of focal attention requires studies that relate fluid intelligence to memory and attention in ways that cannot be accounted for by individual differences in working memory capacity (Shipstead, Harrison, & Engle, submitted). To date few studies have even raised the possibility that such a relationship might exist.

The distinction between complex span and visual arrays performance: current thoughts, limitations, and future directions

Our data and interpretations indicate that complex span and visual arrays tasks largely index the same cognitive mechanisms (see also Shipstead et al., 2014). If so, why do these two types of tasks load on factors that are not perfectly correlated? At present, our own intuition is that the primary difference is in the temporal characteristics of these tasks (see McElree & Dosher, 2001). Visual arrays tasks present information in parallel and require people to maintain the pattern in the absence of physical stimulation. In both of these scenarios, maintenance will be aided by cognitive mechanisms that facilitate access to critical information, even when is not residing in the focus of attention. The predictive differences of complex span and visual arrays performance may thus represent these mechanisms functioning within different task-defined contexts.

For instance, complex span tasks present information in serial order and require integration of information into a list in the face of constant interruption. Retroactive interference is high during the presentation phase of this task. Conversely, visual arrays tasks are free of retroactive interference during item presentation, and instead introduce it at test (Makovski et al., 2010). Therefore, it is virtually guaranteed that every trial of a complex span task requires retrieval of displaced information, thus requiring discrimination against information that was presented on previous trials (Unsworth & Engle, 2006a). In a visual arrays task, interference may occur in a relatively constrained period of time (at test), and thus the need for discrimination against previous trials may be relatively less important, and driven by the task-specific factors (see Lin & Luck, 2012, vs. Shipstead & Engle, 2013, and Souza & Oberauer, 2015).

Of course we are limited in our ability to make these statements conclusively. While the present visual arrays factor is more varied than that of Shipstead, Harrison et al. (2012) and Shipstead, Redick et al. (2012), it remains strictly visuo-spatial in nature. The complex span factor, on the other hand, contains both verbal and visuo-spatial tasks, and is therefore less bound to a particular modality. Thus, these factors differ in more than their temporal characteristics.

In light of this observation it is interesting that the relationships of the complex span and visual arrays factors to fluid intelligence and attention control were quite balanced. Why should the modality-specific tasks be

equally predictive of complex cognition as the cross-modality tasks? This phenomenon may indicate a special relationship between visuo-spatial memory and complex cognition (Süß et al., 2002), or it may represent a bias in testing method (Kane et al., 2004). That is, the correlation may be inflated by the strong spatial components of fluid intelligence and antisaccade tasks (the later strongly defined the attention control factor). More work needs to be conducted relating visual arrays to reasoning and attention in cross-modality manners before conclusive statements can be made.

Despite the observation that visual arrays and complex span performance cannot be cleanly divided into the scope and control of attention, these tasks may nonetheless reflect unique aspects of working memory capacity.

We have already indicated that future studies need to focus on introducing more experimental manipulations to the study of the relationship between working memory, maintenance, attention control, and fluid intelligence. Promising areas of exploration may include a better understanding of the temporal characteristics of these tasks, as well as manipulating the nature of the criterion variables against which complex span and visual arrays have thus far been validated.

Appendix A

Table 5 Upper block of tasks are from Shipstead et al. (2014), lower block of tasks are from Shipstead, Harrison, Trani et al. (submitted for publication)

Day 1	Day 2	Day 3	Day 4
Operation span	**Symmetry span**		
Runing Span	Running Span		
Matrix Task	**Raven**		
VAcolor	**VAcolorS**		
Letter Sets	**Number Series**		
Free Recall	Free Recall		
VAorent	**VAorientS**		
Antisaccade	**Flanker**		
CPA	Split Span		
Digit Span	**Stroop**		
	Beauty Contest		
		Day 3	**Day 4**
Operation span	**Symmetry span**	**Rotation Span**	Paired Learning
Running Span Slow	Runing Span Spatial	Running Span Fast	**Antisaccade-Beep**
Number Series	**Raven**	**Letter Sets**	**Letter Number**
N-Back Words	N-Back Wingdings	N-Back Face	PVT
VAcolor	**VAorient**	**VAorientS**	Paragraph Comp
Garavan 1	Category Shift Task	Number Shift Task	Prospecitve Memory
Antisaccade	**Arrow Flanker**	Trail Making	Prospecitve Memory
N-Back Face	H-Back Words	N-Back Wingdings	Remote Assoicates Task
VerbFluency Animal	VerbFluency C	VerbFluency Occupation	Keep Track
Purdue	**Torrance**	Free Recall	Nonsense Syllogisms
Prospective Memory	Prospective Memory	Cont. Paired Associates	Analogies
Artificial Grammar	Garden Path	**Stroop**	Delayed Paired Learning
Demographics	Cognitive Reflection Test		

note. Reported tasks are bold

Appendix B

Data Set 1

Data Set 2
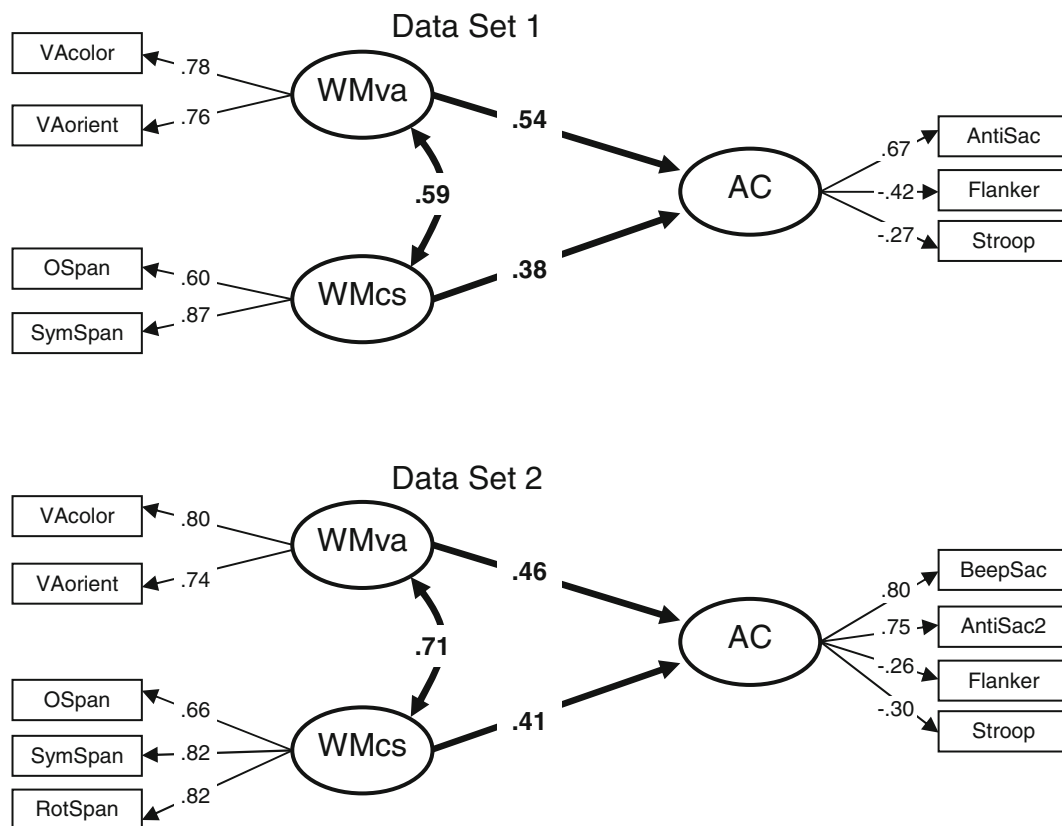
References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317–332.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18,* 622–628.

Awh, E., & Vogel, E. K. (2008). The bouncer in the brain. *Nature Neuroscience, 11,* 5–6.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park: Sage.

Chuderski, A., & Nęcka, E. (2012). The contribution of working memory to fluid reasoning: Capacity, control, or both? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 6,* 1689–1710. doi:10.1037/a002865

Chuderski, A., Taraday, M., Nęcka, E., & Smolén, T. (2012). Storage capacity explains fluid intelligence but executive control does not. *Intelligence, 40,* 278–295.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8,* 331–335.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185. doi:10.1017/S0140525X01003922

Cowan, N., Elliot, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51,* 42–100.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34,* 1754–1768.

Cusack, R., Lehman, M., Veldsman, M., & Mitchell, D. J. (2009). Encoding strategy and not visual working memory capacity correlates with intelligence. *Psychonomic Bulletin & Review, 16,* 641–647.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433. doi:10.3758/BF03214546

Ekstrom, R. B., French, J. W., Harman, M. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton: Educational Testing Service.

Endress, A. D., & Potter, M. C. (2014). Large capacity temporary visual memory. *Journal of Experimental Psychology: General, 143,* 548–565.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23. doi:10.1111/1467-8721.00160

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross

(Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–199). NY: Elsevier.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition, 43,* 226–236.

Fougnie, D., Suchow, J. W., & Alverez. (2012). Variability in the quality of visual working memory. *Nature Communications, 3,* 1229.

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *The Journal of Neuroscience, 29,* 8726–8733. doi:10.1523/JNEUROSCI. 2145-09.2009

Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science, 22,* 361–368. doi:10.1177/0956797611398493

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review, 17,* 673–679.

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-relevant knowledge and experience. *Current Directions in Psychological Science, 20,* 275–279.

Hartshorne, J. K. (2008). Visual working memory capacity and proactive interference. *PLoS ONE, 3*(7), e2716. doi:10.1371/journal.pone.0002716

Healey, M. K., & Miyake, A. (2009). The role of attention during retrieval in working memory span: A dual-task study. *Quarterly Journal of Experimental Psychology, 62,* 733–745.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2013). *Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity.* doi:10.3758/s13428-013-0411-5

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 645–662.

Kane, M. J., Bleckley, K. M., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130,* 169–183.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). NY: Oxford University Press.

Kane, M. J., & Engle, R. W. (2003). 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, 132,* 47–70.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189–217. doi:10.1037/0096-3445.133.2.189

Kim, S. Y., Kim, M. S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences USA, 102,* 16524–16529.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Konstantinou, N., & Lavie, N. (2013). Dissociable roles of different types of working memory load in visual detection. *Journal of Experimental Psychology Human Perception and Performance, 39,* 919–924.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Science, 9,* 75–82.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133,* 339–354.

Lin, P., & Luck, S. J. (2012). Proactive interference does not meaningfully distort visual working memory capacity estimates in the canonical change detection task. *Frontiers in Psychology, 3.*

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17,* 347–356.

Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory & Cognition, 36,* 43–52.

Makovski, T., Watson, L. M., Koutstaal, W., & Jiang, Y. V. (2010). Method matters: Systematic effects of testing procedure on visual working memory sensitivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1466–1479.

Mall, J. T., Morey, C. C., Wolff, M. J., & Lehnert, F. (2014). Visual selective attention is equally functional for individuals with low and high working memory capacity: Evidence from accuracy and eye movements. *Attention, Perception, & Psychophysics*. doi:10.3758/s13414-013-0610-2

McElree, B., & Dosher, B. A. (2001). The focus of attention across space and across time. *Behavioral and Brain Sciences, 24,* 129–130.

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience, 11,* 103–107. doi:10.1038/nn2024

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review, 11,* 296–301.

Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 703–713.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). New York: Oxford University Press.

Raven, J. C. (1990). *Advanced progressive matrices*. Oxford: Oxford Psychological Press.

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and Attention Network Test performance. *Applied Cognitive Psychology, 20,* 713–721.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123,* 374–393. doi:10.1037/0096-3445.123.4.374

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-011-0055-3

Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 277–289. doi:10.1037/a0028467

Shipstead, Z., Harrison, T. L., Trani, A. N., Redick, T. S., Sloan, P., Hicks, K., Bunting, M. F., Draheim, C., & Engle, R. W. (submitted). The Unity and Diversity of Working Memory Capacity and Executive Functions: Their Relationship to General Fluid Intelligence. *manuscript submitted for publication*.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2012). Working memory capacity and visual attention: Top-down and Bottom-up guidance. *Quarterly Journal of Experimental Psychology, 65,* 401–407.

Shipstead, Z., Harrison, T. L., & Engle (submitted). Working memory capacity and fluid intelligence: Maintenance and disengagement. *manuscript submitted for publication*.

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language, 72,* 116–141. doi:10.1016/j.jml.2014.01.004

Shipstead, Z., Redick, T. S., Hicks, K. L., & Engle, R. W. (2012). The scope and control of attention as separate aspects of working memory. *Memory, 20,* 608–628. doi:10.1080/09658211.2012.691519

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology Human Perception and Performance, 31,* 248–261.

Soto, D., Hodsoll, J. P., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12,* 342–348.

Souza, A. S., & Oberauer, K. (2015). Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. *Psychological Bulletin & Review, 22,* 156–162.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze. (2002). Working-memory capacity explains reasoning ability - and a little bit more. *Intelligence, 30,* 261–288. doi:10.1016/S0160-2896(01)00100-3

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154. doi:10.1016/0749-596X(89)90040-5

Unsworth, N. (2009). Examining variation in working memory capacity and retrieval in cued recall. *Memory, 17,* 386–396.

Unsworth, N., & Engle, R. W. (2006a). A temporal-contextual retrieval account of complex span: An analysis of errors. *Journal of Memory and Language, 54,* 346–362.

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

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71,* 1–26. doi:10.1016/j.cogpsych.2014.01.003

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498–505. doi:10.3758/BF03192720

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009a). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17,* 635–654.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1302–1321. doi:10.1037/0278-7393.30.6.1302

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention, Memory, or Both? A direct test of the dual-component model. *Journal of Memory and Language, 62,* 392–406. doi:10.1016/j.jml.2010.02.001

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2009b). Examining the relations among working memory capacity, attention control, and fluid intelligence from a dual-component framework. *Psychology Science Quarterly, 4,* 388–402.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438,* 500–503.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92–114.