

The scope and control of attention: Sources of variance in working memory capacity

Michael Chow · Andrew R. A. Conway

© Psychonomic Society, Inc. 2015

Abstract Working memory capacity is a strong positive predictor of many cognitive abilities, across various domains. The pattern of positive correlations across domains has been interpreted as evidence for a unitary source of inter-individual differences in behavior. However, recent work suggests that there are multiple sources of variance contributing to working memory capacity. The current study ($N = 71$) investigates individual differences in the scope and control of attention, in addition to the number and resolution of items maintained in working memory. Latent variable analyses indicate that the scope and control of attention reflect independent sources of variance and each account for unique variance in general intelligence. Also, estimates of the number of items maintained in working memory are consistent across tasks and related to general intelligence whereas estimates of resolution are task-dependent and not predictive of intelligence. These results provide insight into the structure of working memory, as well as intelligence, and raise new questions about the distinction between number and resolution in visual short-term memory.

Keywords Working memory · Working memory capacity · Individual differences · Visual arrays · Complex span · Intelligence

Working memory is a limited capacity system responsible for the active maintenance of information, as well as its

retrieval from long-term memory. Recent work has emphasized the distinction between two components of working memory: the *scope* and *control* of attention (Cowan et al., 2005). The scope of attention refers to the amount of information that can be actively maintained at a given time, whereas the control of attention refers to the ability to focus on relevant information, and away from irrelevant information.

Working memory is an extremely active area of research in psychology and neuroscience. One motivating factor behind this research effort is the strength of correlations between measures of working memory capacity and measures of higher cognitive function, such as reading comprehension, spatial reasoning, and fluid intelligence (Dane-man & Carpenter, 1980; Kane et al., 2004; Kyllonen & Christal, 1990). While research in this area has traditionally focused on the relationship between the control of attention and higher cognitive function, several recent studies have linked the scope of attention to higher cognitive function as well (Cowan et al., 2005; Fukuda, Vogel, Mayr, & Awh, 2010; Shipstead, Redick, Hicks, & Engle, 2012; Unsworth, Fukuda, Awh, & Vogel, 2014). Moreover, the scope of attention can be fractionated into two components: the number of items that can be maintained and the resolution, or acuity, of those items. Importantly, estimates of number, but not resolution, have been found to correlate with measures of spatial reasoning (Fukuda et al., 2010). This apparent dissociation is important because working memory tasks have been accused of lacking discriminant validity, as they tend to correlate with a broad range of abilities (Ackerman, Beier, & Boyle, 2005; but see Kane, Hambrick, & Conway, 2005). Through examining the scope of attention, the control of attention, and ways in which both facets of working memory are related, this dissociation within the scope of attention

M. Chow · A. R. A. Conway (✉)
Department of Psychology, Princeton University,
Princeton, NJ 08544, USA
e-mail: aconway@princeton.edu

may be leveraged as a powerful tool in understanding the working memory system.

Scope of attention

An influential theory of working memory is the Embedded Process Model (Cowan, 1999). According to this framework, capacity is constrained by the “focus of attention”, which actively maintains items in memory and insulates them from interference and forgetting. In order to assess the number of items that can be simultaneously maintained in the focus of attention, several tasks have been developed. One of the most common procedures is the “visual array” task. For example, Luck and Vogel (1997) used a visual array change-detection procedure, in which participants were required to maintain an array of simple objects in memory, and then indicate whether a probe array was the same or different from the one seen previously. Accuracy was near ceiling when the number of objects in the array was three or less but with four or more objects in the array, accuracy declined. It is also important to note that this pattern of accuracy in detecting changes was the same even when objects increased in complexity. Moreover, by extending a simple model of capacity by Pashler (1988), which takes into account guessing rates, Cowan (2001) demonstrated that capacity estimates tend to be around four items. Importantly, the fact that many other types of visual array tasks, such as those which display only a single item probe, or require the participant to estimate properties of a display item, produce similar capacity estimates and provides strong evidence that the focus of attention can maintain approximately 4 items (Cowan, 2001).

There is also substantial evidence that the number of items held in memory can be distinguished from the resolution, or quality of representation, for those items (Ester, Anderson, Serences, & Awh, 2013; Zhang & Luck, 2008, 2011). Resolution in change detection tasks is often operationalized by increased similarity between probe and memory items, under the rationale that successful recognition requires more detailed memory of the probed item. Using the change detection paradigm, Xu & Chun (2005) found that the intra-parietal sulcus could be functionally dissociated into two distinct regions, one that tracks the number of items regardless of complexity, and another that tracks the number of relevant features. In addition, Awh, Barton, and Vogel (2007) demonstrated that although measures of number and resolution in a change detection task were reliable across different types of stimuli, they were uncorrelated with each other. Finally, Fukuda et al. (2010) compared number, resolution, and spatial reasoning measures using structural equation modeling, and found that only number and spatial reasoning covaried. The correlation

between resolution and spatial reasoning was near zero. In fact, by deriving a 95 % confidence interval from their reported standard errors, it is clear that if a non-trivial correlation exists between resolution and spatial reasoning, it is likely to be fairly weak (r between -0.05 and 0.25).

However, unlike the convergent evidence for a fixed capacity limit across a wide range of tasks, evidence for a dissociation between number and resolution comes primarily from studies that employed only one visual array task. While illustrating this dissociation using a single task is an important step, there is a risk that the effect is driven by task-specific factors. Whereas change detection tasks manipulate the degree of change on a few, discrete levels, more recent tasks have required participants to estimate some property of an item (e.g., color) in a continuous manner (Wilken & Ma, 2004; Zhang & Luck, 2008). Consensus among a range of tasks used to operationalize these constructs would allow for a stronger argument for dissociation. Fortunately, many of the same tasks used to assess capacity limits can be adapted to provide measures of resolution as well.

Control of attention

While the scope of attention refers to the number and resolution of representations that can be held simultaneously in an active state, the control of attention refers to processes that allow for the maintenance of task-relevant representations in the face of distraction (Kane, Bleckley, Conway, & Engle, 2001), as well as selective retrieval from memory (Unsworth & Engle, 2006a). To this end, the control of attention has largely been operationalized through tasks that interject processing components between memoranda, under the rationale that participants must either maintain representations for memoranda throughout the distracting periods, or retrieve them afterward. Such tasks are typically referred to as complex span tasks.

Complex span task scores correlate with performance on tasks that require the top-down guidance of attention. For example, dichotic listening (Conway, Cowan, Bunting, 2001), Stroop (Kane & Engle, 2003), Ericson flanker (Heitz & Engle, 2007), and anti-saccade tasks (Kane et al., 2001; Unsworth, Schrock, & Engle, 2004) all demonstrate relationships to complex span task performance.

In addition, there is mounting evidence that complex span tasks rely on retrieval from episodic memory (Spillers & Unsworth, 2011; Unsworth & Engle, 2006a, b; for a more general overview, see Postle, 2007). One source of evidence comes from the pattern of errors committed on complex span tasks. Consistent with the use of temporal-contextual cues to guide memory search and retrieval, low working memory participants are more likely to omit earlier

list items, allow prior list intrusions for the first position recalled, and have a broader transposition gradient (Unsworth & Engle, 2006a). One explanation for this is that they employ noisier memory cues. Further evidence of this is that they show a reduced tendency to recall words in correct temporal order during delayed free recall, suggesting that recalled words serve more poorly as cues for further recall (Spillers & Unsworth, 2011). Furthermore, memory retrieval during complex span tasks is associated with increased medial temporal lobe activity (Chein, Moore, & Conway, 2011; Faraco et al., 2011). As the medial temporal lobe is thought to be involved with the binding of material in long-term memory (McClelland, McNaughton, & O'Reilly, 1995; O'Reilly, Bhattacharyya, Howard, & Ketz, 2011), these tasks may rely on temporary storage and selective retrieval from episodic memory rather than (or in addition to) active maintenance.

Importantly, convergent evidence that complex span tasks are associated with top-down guidance of attention and retrieval from episodic memory comes from the use of processing components ranging from solving basic arithmetic to making symmetry judgments, and memoranda ranging from letters to spatial locations. Thus, the effects appear to be domain general (Kane et al., 2004).

Comparing scope and control tasks

While much research has focused on either the scope or control of attention individually, there is good reason to believe that they tap largely overlapping constructs. Specifically, Cowan et al. (2005) argued that the critical factor for a valid working memory task is its ability to prevent strategic grouping and rehearsal. Under this line of reasoning, visual array tasks prevent rehearsal by using stimuli that are difficult to verbalize, while complex span tasks prevent rehearsal by interjecting a processing task that occupies attention during intervals between memoranda (Cowan, Saults, & Morey, 2006). Consistent with this reasoning, even complex span tasks with a rudimentary processing component correlate with intelligence (Lépine, Barrouillet, & Camos, 2005). Moreover, for children who are too young to rehearse or group items, simple digit span tasks correlate with intelligence as well (Cowan et al., 2005).

By comparing the scope and control of attention to measures of higher cognitive function, Cowan et al. (2005) were able to assess whether each component contributed unique variance to higher cognitive function. The scope of attention was measured using a wide range of tasks, including visual arrays, which contained only a storage component, but were designed to reduce strategic grouping and rehearsal. The control of attention was assessed

using two complex span tasks. While the complex span tasks explained more variance in higher cognitive function, this variance was task-specific. That is, the additional variance was particular to the individual tasks. Cowan et al. (2005) concluded that disrupting rehearsal in complex span allows a purer estimate of the number of items in memory, implicating these tasks as measures of the scope of attention.

However, Shipstead et al. (2012) highlight an issue that arises from claiming the scope of attention is necessary for explaining the relationship between complex span task performance and higher cognitive function. Namely, for short list lengths, in which the amount of information to be retained falls below capacity estimates for most participants, ceiling effects in measuring capacity should reduce the correlation between complex span and high-order cognition. However, the correlation between complex span and general intelligence remains equally strong across a wide range of list lengths (Bailey, Dunlosky, & Kane, 2011; Salthouse & Pink, 2008; Unsworth & Engle, 2006b). Through reexamining two datasets that contained a visual array task in addition to complex span and general intelligence measures, they concluded that the scope and control of attention represent distinct but largely overlapping constructs. Moreover, the variance shared between visual array and general intelligence measures was largely explained by complex span. However, one limitation they mention is the use of a single visual array task, which is likely to induce a large degree of task-specific variance.

In spite of the strong evidence for overlap between the scope and control of attention, comparisons between the two constructs have yet to address the role of resolution. If resolution is uncorrelated with measures representing the control of attention as well, it represents a fairly astounding finding because over a century of psychometric research suggests that two similar cognitive ability measures should reveal a strong positive correlation, or at the very least, a weak positive correlation. It is extremely rare to find completely uncorrelated variables when testing cognitive ability (Ackerman et al., 2005).

The current study

The present study extends recent work (Fukuda et al., 2010; Shipstead et al., 2012) by administering multiple types of visual array tasks to represent the scope of attention and by administering a broad range of reasoning tasks to provide an accurate assessment of general intelligence. By using multiple measures to reflect the number and resolution of items in the scope of attention, we can reduce the proportion of task-specific variance in that factor, allowing a stronger test of whether it is distinct from the control of attention, and a

more clear view of its shared variance with general intelligence (for an example using many tasks to examine number, but not resolution, see Unsworth et al., 2014). Moreover, assessing and validating resolution from those tasks allows for the opportunity to test whether orthogonality between measures of number and resolution is a general property of visual short-term memory, or a task-specific phenomenon. In addition, it allows for novel comparisons between resolution and the control of attention. If resolution is orthogonal to the control of attention as well, then it could serve as a starting point for development of analogous measures in tasks designed to measure the control of attention.

Method

Participants

Participants ($N = 71$) were recruited from Princeton University and the surrounding community. Students recruited from the Psychology Department participated in exchange for course credit ($n = 55$). Students and community members recruited from the University participant pool ($n = 16$) were compensated \$12/h for their participation.

Procedure

Each participant completed a battery of tasks designed to measure working memory capacity and general intelligence. The tasks were completed in two sessions that lasted approximately two hours each. Breaks were allowed between tasks. General intelligence measures were split evenly between the two sessions. Complex span tasks were administered in the first session, while visual array tasks were administered in the second session. All participants completed general intelligence and working memory tasks in the same, alternating order. In the first session, tasks were administered to groups of up to six participants. In the second session, tasks were administered individually.

Visual array (VA) tasks

Change detection The change detection procedure was based on a task previously developed by Fukuda et al. (2010). For each trial, participants viewed an array of rectangles and ovals. Each shape contained one of two different internal patterns. Participants viewed each array for 500 ms, and after a blank interval of 1,000 ms, were shown a single probe shape with a pattern inside. They then responded as to whether the probe shape was identical to the one they had viewed in the visual array at that location. The probe could either be identical (no change), a different shape (big change), or the same shape but with a different

internal pattern (small change). Arrays contained either 4 or 8 items.

Within each block, half of the trials were no-change trials. For each change trial, there was a 2/3 chance it would be a big change, and a 1/3 chance it would be a small change. Sixteen trials for each set size were interleaved randomly in each block. The total task consisted of nine blocks, and four practice trials at each set size, yielding 296 trials total.

From large change conditions, the number of items in memory was estimated using the formula given by Cowan (2001): $k = \text{set_size} * (\text{correct_hits} - \text{false_alarms})$, where k is the total items in memory, set size is the size of the initial array, correct hits is the proportion of correct change trials, and false alarms are the proportion of incorrect no-change trials. From small change conditions, the resolution of items in memory was calculated in the same way. This was done for each set size. Note that the same no-change trials were used in calculating number and resolution. This is consistent with estimation methods in previous studies (Awh et al., 2007; Fukuda et al., 2010).

Color estimation The color estimation procedure was based on a task previously developed by Zhang and Luck (2008). On each trial, participants viewed an array of 4 or 6 colored squares for 100 ms. Following a 1,000 ms interval, the squares were presented again without color, and a single probe square was outlined in bold. Participants estimated the original color of the probe by selecting from a continuous color wheel, which surrounded the array. As in Zhang and Luck (2008), for each set size, maximum likelihood estimation over a mixture of a von Mises and uniform distribution were used to determine the number and resolution of items simultaneously. Number was represented by the weight given to the von Mises distribution, while resolution was estimated by its dispersion parameter, with greater dispersion parameters representing poor resolution.

Complex span (CS) tasks

Three automated complex span tasks were administered (Unsworth, Heitz, Schrock, & Engle, 2005; see also Redick et al., 2012). On each trial of these tasks, participants alternated between performing a secondary task and viewing a to-be-remembered (TBR) item. After a number of rounds, participants reported as many TBR items as possible, in the same order they were presented, by selecting the items from a grid. A point was awarded for each TBR item recalled in correct serial order, which is often referred to as partial credit load scoring Conway et al. (2005). Participants were instructed to click a button labeled “blank” to skip a serial position.

Operation span The operation span (OSPAN) consisted of solving basic arithmetic, followed by a TBR letter. Set sizes ranged from 3 to 7 items. Each set size was completed 3 times.

Reading span The reading span (RSPAN) consisted of indicating whether a sentence was grammatical, followed by a TBR letter. Set sizes ranged from 3 to 7 items. Each set size was completed 3 times.

Symmetry span The symmetry span (SSPAN) task consisted of solving for whether an 8×8 grid of black and white squares was symmetrical, followed by a TBR blue square presented on a 4×4 grid. Set sizes ranged from 2 to 6 items. Each set size was completed 3 times.

Intelligence tasks

All measures of intelligence were administered on paper. Before each task, an experimenter reviewed the instructions aloud, went through practice problems, and asked participants whether they had any questions. For each task, one point was awarded for each correct response.

Raven's advanced progressive matrices, set II In this task, a 3×3 matrix has 8 images in it, while the bottom-right cell is missing in each problem. Participants indicate which of eight options would complete the matrix by looking at the relations among the images across the rows and down the columns. After completing the first two problems from Set I as practice, participants were given the 18 odd numbered problems from Set II, and had 15 min to complete the task.

Cattell's culture fair task This task consisted of four subtests. In the first, a row of three boxes and an empty box were shown, and participants chose which option best completes a pattern shown in the first three boxes. In the second, five boxes were shown and participants indicate which two boxes do not belong to the pattern made by the other three boxes. In the third, images fill a 2×2 or 3×3 matrix of boxes, but the bottom-right box was empty. Participants chose, from a number of options, the box that best completed the pattern. Finally, the last test consisted of an image containing a number of shapes and a dot. The goal of this test was to choose from different configurations of shapes, the one in which a dot could be placed in a similar location relative to the shapes (e.g., encompassed by a triangle, but not by a circle). In some instances, the images contained two dots, and participants were to follow a similar rule in choosing responses. The number of items for each subtest were 13, 14, 13, and 10, respectively. Scores for each subtest were summed to create one overall score.

DAT space relations test For each item, flat shapes were presented along with five printed 3-dimensional depictions of what each shape might look like when folded along its edges. Participants chose the 3-dimensional depiction that could be made from the flat shape. This test consisted of 18 items.

ETS surface development For each item, a flat shape was presented along with an accurate 3-dimensional depiction of that shape folded along its edges and rotated. Lines on the flat shape were numbered, while lines on the folded shape had letters next to them. Participants indicated which numbers on the flat shape corresponded to letters on the folded shape. This test consisted of 5 shapes with 5 items each, yielding 25 items total.

AFQT reading comprehension For each item, participants read short paragraphs and chose from five options the one that best completed each paragraph. This test consisted of 15 items.

AFQT analogies For each item, participants were presented with an incomplete analogy (e.g., QUART is to LITER as INCH is to) and chose from five options the one that best completed each analogy. This test consisted of 18 items.

Results

Data screening

Participants with scores greater than 4 standard deviations on any measure were considered univariate outliers and discarded. From this procedure, 2 participants were removed. In addition, the mvoutlier package in R was used to plot the ordered, robust Mahalanobis distance against the quantiles of the Chi-squared distribution. Participants who deviated largely from the expected quantiles were removed and the plot was reconstructed until there appeared to be no substantial curvature. From this procedure, 2 additional participants were discarded as multivariate outliers.

Statistical procedures

A series of latent variable models (aka confirmatory factor analyses) were conducted. Multiple indices of fit are reported for each model. First, the Chi-square statistic measures the mismatch between the observed and reconstructed covariance matrices. Higher values indicate greater mismatch. However, moderate to large sample sizes may lead to statistically significant differences, even when the discrepancy between covariance matrices is slight. So in addition,

we report the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), non-normed fit index (NNFI), comparative fit index (CFI), and Akaike's information criterion (AIC). Following the suggestion of Hu and Bentler (1999), we consider the following to be evidence of good fit: RMSEA < .08, SRMR < .08, NNFI > .95, CFI > .95. AIC is a measure of parsimony used for model comparisons, with lower AIC indicating a better fit after applying a penalty for free parameters.

All analyses were performed using the R software language (R Core Team 2012). In addition to the base R libraries, custom libraries were used for general data processing (Wickham, 2007, 2011), outlier detection (Filzmoser & Gschwandtner, 2013), and structural equation models (Rosseel 2012).

Descriptive statistics

Descriptive statistics for all measures are provided in Table 1. All meet the acceptable criteria suggested by Kline (2011) for latent variable model analyses (absolute skew < 3; kurtosis < 10). The full correlation matrix is provided in Table 2.

Resolution and number for visual arrays

We tested whether measures for number and resolution in the scope of attention load on distinct factors, and whether resolution is task dependent, by contrasting three models. In the first model, all visual array measures loaded on a single factor. In the second model, measures of resolution and number loaded on separate factors. The third model was similar to the second, except that the resolution factor was split into two, so that each task loaded onto a separate resolution factor.

Fit indices are provided in Table 3. All models failed to meet any of the criteria for a good fit to the data. In order to address whether the poor fit was because of task-specific covariance among measurements, new models were constructed, which allowed residuals to covary. Specifically, for each task and each measure (either number or resolution), the residuals were allowed to covary across the two set sizes. However, all fit indices for these models also failed to meet the criteria for a good fit.

All three of our theoretically motivated models failed to fit the data. In an attempt to find a model that could fit the data, we considered a four-factor model, consisting of two factors for number and two factors for resolution. This

Table 1 Descriptive statistics

	mean	sd	min	max	skew	kurtosis
s span	31.15	7.67	13.00	42.00	−0.47	−0.81
r span	62.12	10.48	24.00	75.00	−1.58	3.20
o span	63.54	9.61	36.00	75.00	−1.05	0.61
cd_4_num	2.96	0.57	1.46	3.89	−0.86	−0.03
cd_8_num	3.65	1.45	0.31	6.59	−0.05	−0.92
cd_4_res	1.59	0.75	−0.25	3.52	−0.09	−0.27
cd_8_res	1.37	1.18	−1.38	3.64	−0.04	−0.44
color_4_num	66.96	17.38	21.82	92.30	−0.72	−0.09
color_6_num	47.91	15.14	5.40	78.84	−0.43	0.23
color_4_res	19.97	3.90	8.56	33.15	0.35	1.43
color_6_res	21.99	7.38	5.73	52.99	1.01	3.16
ravens	13.70	2.36	8.00	18.00	−0.27	−0.54
cft	31.12	3.82	21.00	39.00	0.01	−0.38
surfdev	20.43	4.93	6.00	25.00	−1.10	0.33
spacerel	13.40	2.94	6.00	18.00	−0.52	−0.36
readcomp	11.69	2.05	6.00	14.00	−0.74	−0.25
analogy	13.19	2.07	8.00	17.00	−0.40	−0.34

rspan = reading span; sspan = symmetry span; ospan = operation span; cd-4-num = change detection (number), set size 4; cd-8-num = change detection (number), set size 8; cd-4-res = change detection (resolution); cd-8-res = change detection (resolution); color-4-num = color estimation (number), set size 4; color-6-num = color estimation (number), set size 6; color-4-res = color estimation (resolution), set size 4; color-6-res = color estimation (resolution), set size 6; ravens = Raven's Advanced progressive matrices; cft = Cattell's culture-fair test; surfdev = surface development; spacerel = space relations; analogy = analogy battery; readcomp = reading comprehension

Table 2 Correlations among variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
sspan	—																
rspan	.50***	—															
ospan	.22	.57***	—														
cd_4_num	.32**	.34**	.13	—													
cd_8_num	.52***	.38**	.24	.66***	—												
cd_4_res	.39***	.20	.19	.45***	.45***	—											
cd_8_res	.17	-.01	.17	.26*	.43***	.58***	—										
color_4_num	.26*	.21	.03	.58***	.43***	.43***	.27*	—									
color_6_num	.15	.13	-.05	.50***	.38**	.38**	.17	.77***	—								
color_4_res	-.25*	-.18	-.11	-.07	.02	-.05	.02	.15	.14	—							
color_6_res	-.16	.00	-.07	.06	.03	-.24	-.23	.12	.41***	.47***	—						
ravens	.43***	.26*	.01	.41***	.61***	.30*	.15	.29*	.34**	-.11	.03	—					
cft	.40***	.23	.06	.34**	.42***	.27*	.17	.22	.08	-.11	-.19	.37**	—				
surfdev	.58***	.46***	.26*	.38**	.51***	.37**	.14	.22	.29*	-.04	.06	.54***	.49***	—			
spacerel	.47***	.37**	.22	.50***	.49***	.39**	.21	.39**	.39**	.11	.06	.58***	.37**	.67***	—		
readcomp	.25*	.35**	.27*	.11	.20	.11	.08	.09	.11	-.02	.01	.22	.19	.38**	.28*	—	
analogy	.22	.38**	.31*	.26*	.31*	.17	.02	.07	.02	-.15	-.08	.40***	.33**	.44***	.35**	.27*	—

rspan = reading span; sspan = symmetry span; ospan = operation span; cd-4-num = change detection (number), set size 4; cd-8-num = change detection (number), set size 8; cd-4-res = change detection (resolution); cd-8-res = change detection (resolution); color-4-num = color estimation (number), set size 4; color-6-num = color estimation (number), set size 6; color-4-res = color estimation (resolution), set size 4; color-6-res = color estimation (resolution), set size 6; ravens = Raven's Advanced Progressive Matrices; cft = Cattell's culture-fair test; spacerel = space relations; analogy = analogy battery; readcomp = reading comprehension

Table 3 Fit indices for all visual array tasks

	model	chisq	df	pvalue	rmsea	srmr	nnfi	cfi	aic
1	vstm.1f	108.91	20.00	0.00	0.26	0.15	0.42	0.59	1410.44
2	vstm.2f	87.54	19.00	0.00	0.23	0.15	0.53	0.68	1391.07
3	vstm.3f.2res	58.17	17.00	0.00	0.19	0.09	0.69	0.81	1365.71
4	vstm.4f	44.51	18.00	0.00	0.15	0.09	0.81	0.88	1350.04

rmsea = root mean square error of approximation; srmr = standardised root mean square residual; nnfi = non-normed fit index; cfi = comparative fit index; aic = Akaike's information criterion

model provided a better fit than the other models but still did not satisfy our criteria for adequate fit (see Table 3).

An examination of the correlation matrix shows that measures of resolution are correlated within each task (change detection, $r = .58$, $p < .001$; color estimation, $r = .47$, $p < .001$). However, the correlations between number and resolution *across* tasks is complicated. For the change detection task the correlations between number and resolution are positive at small and large set sizes. In contrast, for color estimation, the correlations between number and resolution vary as a function of set size. Different patterns of correlation across set sizes within the visual array tasks are not predicted by any of the models, which helps to explain the relatively poor fit of our three theoretically motivated models.

One vs. two factors of working memory capacity

In order to assess whether visual array and complex span tasks are better represented by distinct components, we compared a single-factor model to a two-factor model. In the single-factor model, visual array and complex span tasks loaded on the same factor. In the two-factor model, visual array tasks loaded on one factor, while complex span tasks loaded on another. Consistent with previous modeling efforts using automated span tasks (Shipstead et al., 2012), residuals for OSPAN and RSPAN were allowed to covary in each of the models.

Fit indices are provided in Table 4. A direct comparison of the models favored the two-factor model ($\chi^2 = 9.47$, $p < 0.01$). All fit statistics for the single-factor model failed to meet the criteria listed in the Statistical Procedure

Table 4 Fit indices for working memory tasks

	model	chisq	df	pvalue	rmsea	srmr	nnfi	cfi	aic
1	1.factor	22.67	12.00	0.03	0.12	0.08	0.90	0.94	1178.05
2	2.factor	13.94	11.00	0.24	0.06	0.06	0.97	0.98	1171.33

rmsea = root mean square error of approximation; srmr = standardised root mean square residual; nnfi = non-normed fit index; cfi = comparative fit index; aic = Akaike's information criterion

section, indicating a poor overall fit to the data. However, all fit statistics for the two-factor model met or exceeded the same criteria, indicating a good overall fit. The two-factor model is displayed in Fig. 1. Consistent with Shipstead et al. (2012), the latent variable for visual array tasks is labeled VA, while the latent variable for complex span tasks is labeled CS.

Model of intelligence and g

The battery of intelligence tests consisted of verbal, spatial, and fluid reasoning tasks. The battery consisted of only 6 tests but we tapped different domains in order to estimate a general factor of intelligence. Indeed, a one-factor model of the 6 intelligence tests proved to be sufficient. The model is illustrated in Fig. 2. All fit statistics satisfied the criteria listed in the Statistical Procedure section, indicating a good overall fit of the model.

The relationship of VA and CS to g

The models tested thus far suggest that (a) VA and CS are separate but related factors; and (b) a strong general factor emerged from the battery of intelligence tests we administered. Based on these results, we tested the hypothesis that VA and CS each account for unique variance in g by computing factor scores for VA, CS, and g (based on the models reported above) and then conducting a multiple regression analysis with g as the outcome variable and VA and CS as the predictors. A summary of the regression analysis is reported in Table 5. Our prediction that VA and CS account for independent sources of variance in g was supported. The regression coefficients for both VA and CS were significant (for both, $p < .05$) and together VA and CS accounted for 85 % of the variance in g .

Finally, to be consistent with prior work, we considered models that included all manifest variables from the entire study. Based on Shipstead et al. (2012), we tested three models. The first model predicts that VA is the primary source of variance in g . The second model predicts that CS is the primary source of variance in g . The third model, which we prefer, and which is consistent with our regression analysis,

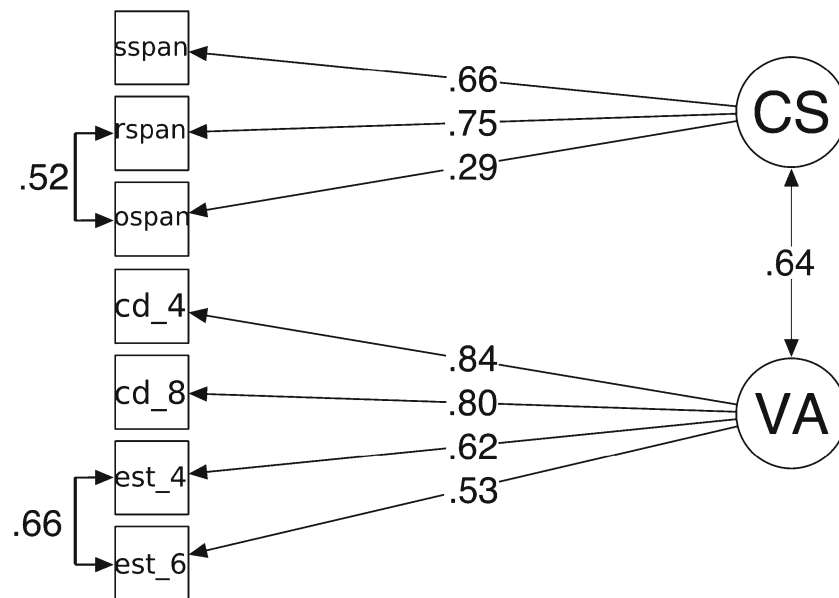


Fig. 1 Two factor model for working memory tasks. est-4 = color estimation, set size 4; est-6 = color estimation, set size 6; cd-4 = change detection, set size 4; cd-8 = change detection, set size 8; rspan = reading span, sspan = symmetry span, ospan = operation span

assumes that both VA and CS account for unique variance in g . The fit statistics for these models are provided in Table 6, and SEM parameter estimates for the third model are presented in Fig. 3. The second and third models are clearly superior to the first but the distinction between model 2 and model 3 is less clear. While model 2 is more parsimonious, we are hesitant to conclude that CS is essential and VA is irrelevant. The above regression analysis, and previous work, suggests that both CS and VA contribute to variability in general intelligence. The lack of discrimination between models 2 and 3 in the current study may be due to the small sample and/or the limited sample of tasks for both CS and VA.

In order to ensure that the current findings are not a result of the large number of measurements, relative to number of participants, the same latent variable models were examined, but with measures that came from the same type of visual short-term memory task (either change detection or color estimation) averaged together. This is known as measurement parceling, and is done to both reduce the number of free parameters, as well as increase the test reliability of each resulting measurement (Little, Cunningham, Shahrar, & Widaman, 2002). Consistent with advice given by

Little et al., (2002) for reducing subfactors to parcels, in order to gain a general, unidimensional measure of g , each pair of tasks were averaged within the spatial, formal inductive, and verbal reasoning domains. Complex span tasks were not parceled, so that overall there were eight measurements. Comparisons among the models with measurement parcels were consistent with the original models. See Table 7 for model fits, and Fig. 4 for SEM estimates of the model with both CS and VA as predictors of intelligence.

Discussion

Relationship between VA and CS

The current results provide further support for the argument that visual array tasks and complex span tasks tap distinct but related constructs, replicating and extending recent studies by Shipstead et al. (2012) and Unsworth et al. (2014). Critically, the use of visual array tasks that are common in research, as well as a broad set of intelligence measures provides convergent evidence for an important distinction between these two task paradigms. Moreover, the latent correlation between VA and CS observed here was intermediate to those obtained by Shipstead et al. (2012) from two separate data sets ($r = .64$ compared to $r = .61$ and $r = .69$). Though the visual array tasks used here differed from Shipstead et al. (2012) in several critical ways, and intelligence measures from three subdomains were used, these findings replicate the same pattern of results from their two-factor working memory models.

Table 5 Summary of Multiple Regression Analysis for g

predictor	β	SE(β)	t	Sig. (p)
VA	.158	.065	2.41	.018
CS	.808	.065	12.32	< .001

$R^2 = .851$. VA = visual array (capacity) tasks; CS = complex span tasks

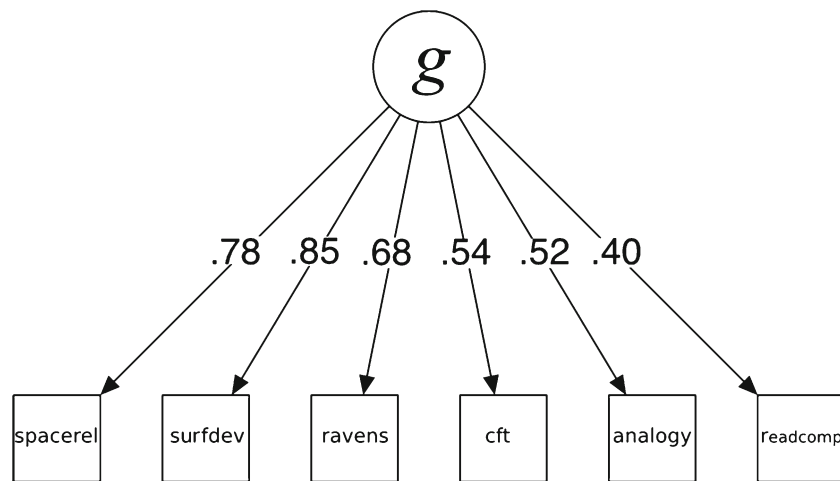


Fig. 2 One factor model of intelligence. ravens = Ravens Advanced progressive matrices; cft = Cattell's culture-fair test; spacerel = space relations; analogy = analogy battery; readcomp = reading comprehension

Moreover, the constructs underlying visual array tasks and complex span tasks each account for unique variance in general intelligence. This finding supports arguments that there are multiple sources of variance underlying intelligence and working memory capacity. This result is inconsistent with unitary source models of general cognitive ability. The idea that there are multiple sources of variance underlying general cognitive abilities such as working memory capacity and intelligence has a long history (e.g., Thurstone, 1931, 1938) and is consistent with multi-component models of working memory (Baddeley & Hitch, 1974), as well as prior investigations of individual differences in working memory capacity that have posited multiple pools of resources (Daneman & Tardiff, 1987; Shah & Miyake, 1996). Indeed, the current results are consistent with the notion that there are multiple sources of variation across both task domains and cognitive processes (Conway & Kovacks, 2013; Logie, 2011; Oberauer, Süß, Wilhelm, & Wittmann, 2008).

Orthogonality of number and resolution

Contrary to previous studies (Awh et al., 2007; Fukuda et al., 2010; Scolari, Vogel, & Awh, 2008), which did not

find evidence of correlation between measures of number and resolution, our change detection and color estimation tasks showed moderately strong correlations between resolution and number, though the color estimation task did so only at set size 6. If accurate, these results go beyond suggesting that the orthogonality of number and resolution may be task specific, and question the validity of the number and resolution constructs, as they were measured in these tasks. Given that the correlations appear selective to different set sizes, it appears that the relationship between number and resolution may not be as straightforward as previous accounts have claimed. While previous findings suggest that resolution and number are uncorrelated in change detection tasks, some degree of caution seems warranted. That number and resolution are correlated in the current change detection task is puzzling, as it was constructed using the stimuli and procedures given by Fukuda et al. (2010). However, there were a few differences between their study and the current procedure, which may help explain the conflicting outcomes.

First, unlike Fukuda et al. (2010), which averaged Cowan's k scores across set sizes, the current study averaged scores across stimuli, before calculating Cowan's k for each set size. Averaging across set sizes is problematic in that capacity estimates at set sizes that do not exceed a participant's capacity are necessarily biased downward (Rouder, Morey, Morey, & Cowan, 2011). However, calculating Cowan's k in the manner of Fukuda et al. (2010) still produced significant correlations between number and resolution measures.

Second, although the stimuli for both versions of the task were the same, the exact instructions differed. Thus, it is possible that a large subset of participants misunderstood task instructions and performed poorly in all aspects of the task. This could induce correlation between all measures.

Table 6 Fit indices for model of intelligence

model	chisq	df	pvalue	rmsea	srmr	nnfi	cfi	aic
1 sem.va	75.86	60.00	0.08	0.06	0.08	0.94	0.95	2178.06
2 sem.cs	69.48	60.00	0.19	0.05	0.07	0.96	0.97	2171.69
3 sem.both	67.57	59.00	0.21	0.05	0.07	0.97	0.97	2171.77

rmsea = root mean square error of approximation; srmr = standardised root mean square residual; nnfi = non-normed fit index; cfi = comparative fit index; aic = Akaike's information criterion

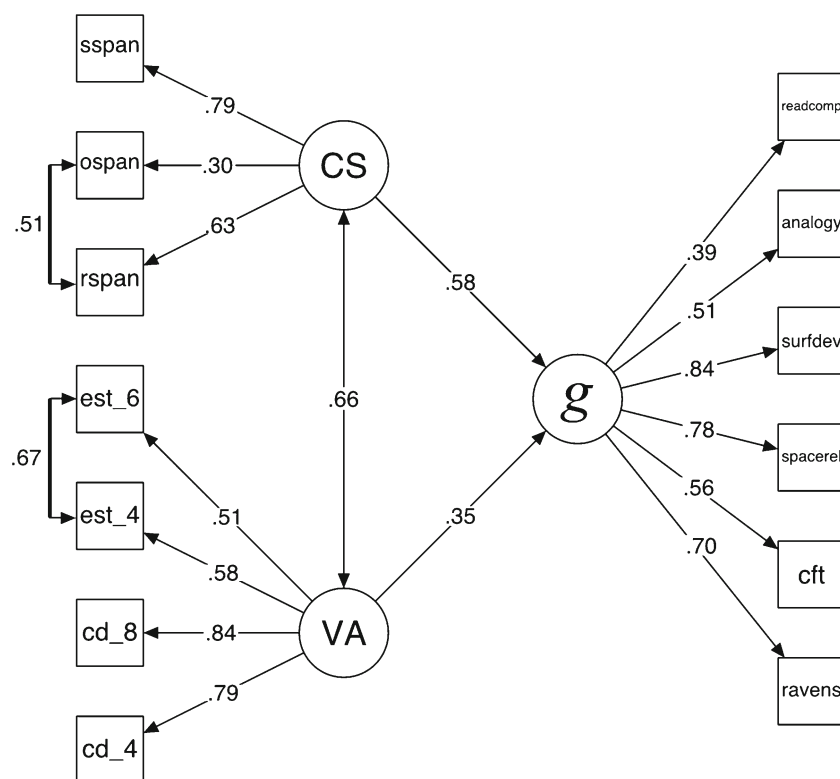


Fig. 3 Full model. est = color estimation; cd = change detection; rspan = reading span; sspan = symmetry span; ospan = operation span; ravens = Raven's Advanced progressive matrices; cft = Cattell's

culture-fair test; spacerel = space relations; analogy = analogy battery; readcomp = reading comprehension

However, the number of capacity estimates that were at or below zero, which indicates performance around chance levels, in this task appears comparable to Fukuda et al. (2010). It is worth noting that in both studies a substantial number of participants fell near or below zero for resolution estimates.

A third possibility, which addresses the large number of low or negative resolution estimates, is that mixing small and large change trials in the same blocks produces capacity estimates which may not correctly control for response bias due to guessing. This is because Cowan's *k* only takes into account response bias in the case where an object (or

relevant feature) is retained in memory or lost completely (Rouder et al., 2011). While information about whether a single feature of an object is present in a probe may be sufficient to detect a change from sample to test, both features of the object are necessary to determine whether there was no change. Thus, while the model takes into account only a single guessing rate, participants may employ a guessing rate based on whether they retained only the probed shape in memory, and a separate guessing rate if they retained neither relevant feature in memory. A participant could strategically reduce overall false alarms at the expense of change detection accuracy for small changes by always indicating "no change" if they only know that a large change has not occurred. This might occur, for example, if the participant found the small change component to be too difficult.

Differential strategy use of this nature could artificially reduce or eliminate correlations between number and resolution, while keeping both reliable across set sizes. Moreover, in reanalyzing the change detection data of Fukuda et al. (2010), a pattern emerges which seems to be consistent with this explanation. Specifically, residuals for the linear models regressing resolution on number appear U-shaped (see Fig. 5), with a cluster of participants who scored near

Table 7 Fit indices for model of intelligence with measure parcels

model	chisq	df	pvalue	rmsea	srmr	nnfi	cfi	aic
1 sem.va	32.34	17.00	0.01	0.12	0.08	0.87	0.92	1365.10
2 sem.cs	24.95	17.00	0.10	0.08	0.07	0.93	0.96	1357.71
3 sem.both	22.63	16.00	0.12	0.08	0.07	0.94	0.97	1357.39

rmsea = root mean square error of approximation; srmr = standardised root mean square residual; nnfi = non-normed fit index; cfi = comparative fit index; aic = Akaike's information criterion

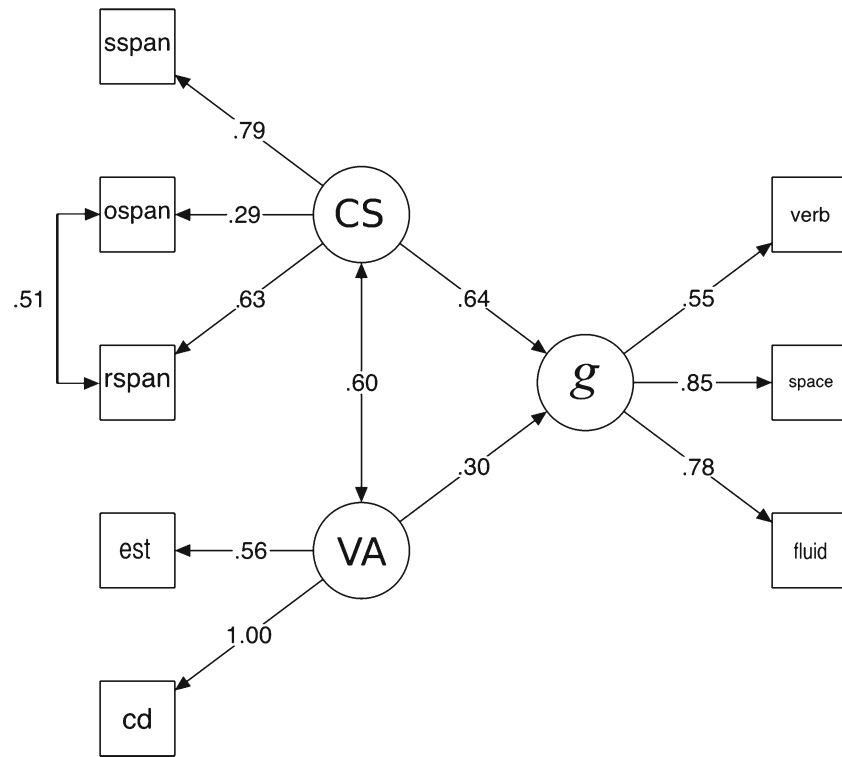


Fig. 4 Full model using measure parcels. est = color estimation; cd = change detection; rspan = reading span; sspan = symmetry span; ospan = operation span; verb = verbal intelligence parcel; space = spatial intelligence parcel; fluid = fluid intelligence parcel

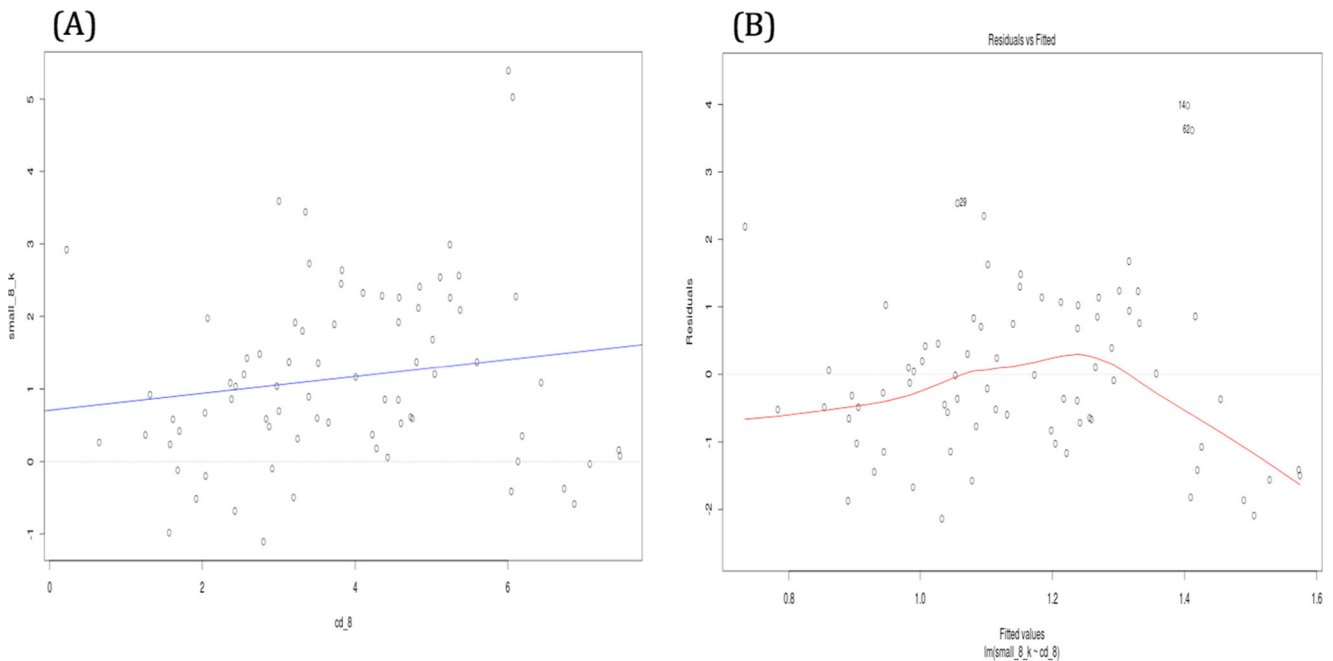


Fig. 5 Resolution (small_change_k) regressed on number (cd_8) for set size 8 of a change detection task from Fukuda et al (2010). Panel **A** shows a scatter plot with the best fitting linear model in blue. Panel **B**

shows the residuals versus fitted values from the linear model. A clear U-shaped pattern can be observed in the residuals

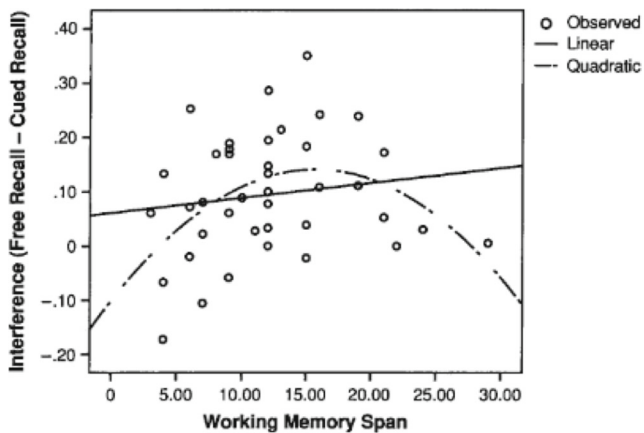


Fig. 6 Plot taken from Cokely et al. (2006) that shows a distinct U-shaped pattern. After questioning participants about their strategy use, they discovered that high performers on complex span tasks had used a strategy which negated the effect of their partial list cueing paradigm. The solid line traces the best fitting linear model, while the dotted line traces the best fitting quadratic model

zero for resolution, but high for number. It is possible that high capacity participants enhanced number estimates by sacrificing performance on small change trials. Intriguingly, a similar issue arose in an experiment by Cokely, Kelley, and Gilchrist (2006), which related complex span performance to a partial list cueing paradigm. Initially, complex span performance was found to be uncorrelated with a partial list cueing effect. However, the authors discovered that high complex span participants were engaging in a strategy that negated the anticipated effect, creating U-shaped residuals (see Fig. 6). Upon controlling for strategy use, a correlation between complex span and the partial list-cueing paradigm emerged.

However, the finding that resolution estimates in the color estimation task were reliable within task, but correlated neither with number nor resolution in the change detection task raises the possibility that the change detection task used in the current study failed to yield valid measures of resolution. However, this explanation fails to account for the correlation between resolution and number at set size 6 in the color estimation task.

Whether the cluster of participants observed in the change detection data of Fukuda et al. (2010) represent a meaningful pattern, or the current study failed to accurately tap into resolution and number estimates is unclear. Further control and replication are necessary to clarify whether resolution and number are correlated. One possibility for controlling the potential contributions of strategy use in the change detection paradigm would be to administer blocks with only one type of change trial in addition to mixed blocks. Blocks with one trial type would allow for an unbiased assessment of resolution and number, as participants

would have little incentive to reduce resolution performance. Moreover, performance on these blocks could be compared with mixed blocks, to assess whether participants make strategic trade-offs.

Finally, there is a long-standing debate over whether performance on the change detection and color estimation tasks is better reflected by the slot-based models used in the current study, where items are remembered in an all-or-nothing fashion, or models in which a continuous memory resource is distributed across items (Bays & Husain, 2008; Cowan & Rouder, 2009; Luck & Vogel, 1997; Rouder et al., 2011; van den Berg, Shin, Chou, George, & Ma, 2012). While previous studies examining the relationship between visual working memory capacity and intelligence (Fukuda et al., 2010; Shipstead et al., 2012; Unsworth et al., 2014) have also assumed that the slot-based models are valid, the failure to find support for these models in our simple latent variable analyses of resolution across tasks raise the possibility that other accounts may be more useful. This is especially important to note when considering the findings of Unsworth (2014), as they (presumably) used measures where capacity and resolution are often estimated simultaneously using slot-based models (as in Zhang & Luck, 2011), but make no mention of resolution estimates. It would be interesting to see whether they find a common factor underlying resolution estimates across tasks. Another worthwhile avenue of research would be to investigate how these alternative models account for individual differences in visual working memory performance. Evidence in favor of models that distribute a continuous memory resource across items was reviewed recently by Ma, Husain, and Bays (2014).

Conclusions

The current study replicates and extends recent findings by Shipstead et al. (2012) suggesting that the scope and control of attention contribute independent sources of variance to working memory capacity and general intelligence. These findings support multiple source theories of capacity and intelligence and are inconsistent with general ability models (cf., Conway & Kovaks, 2013; Logie, 2011; Oberauer et al., 2008). The current study failed to replicate certain aspects of Fukuda et al. (2010) with respect to the relationship between number and resolution measures derived from visual array tasks. Several possibilities, including differential strategy use across studies, might explain these differences. However, it is clear that more work is necessary to leverage individual differences in order to assess the validity of number and resolution as conceptualized by slot-based models of visual short-term memory, as well

as the claim that resolution and number in the scope of attention are orthogonal.

References

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, 131(1), 30.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, 18(7), 622–628.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In B. G. (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Bailey, H., Dunlosky, J., Kane, M. J. (2011). Contribution of strategy use to performance on complex and simple span tasks. *Memory & Cognition*, 39(3), 447–461.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321(5890), 851–854. doi:10.1126/science.1158023
- Chein, J. M., Moore, A. B., & Conway, A. R. A. (2011). Domain-general mechanisms of complex working memory span. *NeuroImage*, 54(1), 550–559.
- Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, 13(6), 991–997.
- Cowan, N. (1999). An embedded-processes model of working memory. *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, 62–101.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–185.
- Cowan, N., Elliott, 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(1), 42–100.
- Cowan, N., & Roudier, J. N. (2009). Comment on “Dynamic shifts of limited working memory resources in human vision”. *Science*, 323(5916), 877c.
- Cowan, N., Saults, J. S., & Morey, C. C. (2006). Development of working memory for verbal-spatial associations. *Journal of Memory and Language*, 55(2), 274–289.
- Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review*, 8(2), 331–335.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and users guide. *Psychonomic Bulletin and Review*, 12(5), 769–786.
- Conway, A. R., & Kovacks, K. (2013). Individual differences in intelligence and working memory: A review of latent variable models. *Psychology of Learning and Motivation*, 58, 233–270.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19(4), 450–466.
- Daneman, M., & Tardif, T. (1987). Working memory and reading skill reexamined. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 491–508). Hillsdale, NJ: Erlbaum.
- Ester, E. F., Anderson, D. E., Serences, J. T., & Awh, E. (2013). A neural measure of precision in visual working memory. *Journal of Cognitive Neuroscience*, 25(5), 754–761. doi:10.1162/jocn.a.00357
- Faraco, C. C., Unsworth, N., Langley, J., Terry, D., Li, K., Zhang, D., ... Miller, L. S. (2011). Complex span tasks and hippocampal recruitment during working memory. *NeuroImage*, 55(2), 773–787.
- Filzmoser, P., & Gschwandtner, M. (2013). *mvoutlier: Multivariate outlier detection based on robust methods*. Retrieved from <http://CRAN.R-project.org/package=mvoutlier>
- 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(5), 673–679.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General*, 136(2), 217.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. doi:10.1080/10705519909540118
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130(2), 169–183.
- 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(1), 47.
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle. *Psychological Bulletin*, 131(1), 66–71.
- 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(2), 189–217.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*, 3rd edn. New York: Guilford.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, 14(4), 389–433. doi:10.1016/S0160-2896(05)80012-1
- Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin & Review*, 12(1), 165–170.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9(2), 151–173.
- Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, 20, 240–245.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–280.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(3), 347–356. doi:10.1038/nn.3655
- McClelland, J. L., McNaughton, B., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457.
- Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence*, 36, 641–652.

- O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2011). Complementary learning systems. *Cognitive Science*, 1, 1–20. doi:[10.1111/j.1551-6709.2011.01214.x](https://doi.org/10.1111/j.1551-6709.2011.01214.x)
- Pashler, H. (1988). Familiarity and visual change detection. *Attention, Perception, & Psychophysics*, 44(4), 369–378.
- Postle, B. R. (2007). Activated long-term memory? The bases of representation in working memory. *The Cognitive Neuroscience of Working Memory*, 333–349.
- R Core Team (2012). R: A language and environment for statistical computing. Vienna, Austria.
- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., ... Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, 28(3), 164.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- 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*, 18(2), 324–330.
- Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review*, 15(2), 364–371.
- Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review*, 15(1), 215–222. doi:[10.3758/PBR.15.1.215](https://doi.org/10.3758/PBR.15.1.215)
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, 125, 4–27.
- 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(6), 608–628.
- Spillers, G. J., & Unsworth, N. (2011). Variation in working memory capacity and temporal-contextual retrieval from episodic memory.
- Thurstone, L. L. (1931). A multiple factor study of vocational interests. *Personnel Journal*, 10, 198–205.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Unsworth, N., & Engle, R. W. (2006a). A temporal-contextual retrieval account of complex span: An analysis of errors. *Journal of Memory and Language*, 54(3), 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(1), 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](https://doi.org/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(3), 498–505.
- 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(6), 1302.
- Van den Berg, R., Shin, H., Chou, W.-C., George, R., Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, 109(22), 8780–8785. doi:[10.1073/pnas.1117465109](https://doi.org/10.1073/pnas.1117465109)
- Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12), 1–20.
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1), 1–29.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4(12), 11–11. doi:[10.1167/4.12.11](https://doi.org/10.1167/4.12.11)
- Xu, Y., & Chun, M. M. (2005). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440(7080), 91–95. doi:[10.1038/nature04262](https://doi.org/10.1038/nature04262)
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. doi:[10.1038/nature06860](https://doi.org/10.1038/nature06860)
- Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, 22(11), 1434–1441. doi:[10.1177/0956797611417006](https://doi.org/10.1177/0956797611417006)