

Associative learning predicts intelligence above and beyond working memory and processing speed

Scott Barry Kaufman^{a,*}, Colin G. DeYoung^b, Jeremy R. Gray^{a,c}, Jamie Brown^d,
Nicholas Mackintosh^d

^a Yale University, Department of Psychology, USA

^b University of Minnesota, Department of Psychology, USA

^c Yale University, Interdepartmental Neuroscience Program, USA

^d University of Cambridge, Department of Experimental Psychology, UK

ARTICLE INFO

Article history:

Received 25 April 2008

Received in revised form 16 February 2009

Accepted 16 March 2009

Available online 9 April 2009

Keywords:

Intelligence

Associative learning

Working memory

Processing speed

Elementary cognitive processes

ABSTRACT

Recent evidence suggests the existence of multiple cognitive mechanisms that support the general cognitive ability factor (*g*). Working memory and processing speed are the two best established candidate mechanisms. Relatively little attention has been given to the possibility that associative learning is an additional mechanism contributing to *g*. The present study tested the hypothesis that associative learning ability, as assessed by psychometrically sound associative learning tasks, would predict variance in *g* above and beyond the variance predicted by working memory capacity and processing speed. This hypothesis was confirmed in a sample of 169 adolescents, using structural equation modeling. Associative learning, working memory, and processing speed all contributed significant unique variance to *g*, indicating not only that multiple elementary cognitive processes underlie intelligence, but also the novel finding that associative learning is one such process.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Over a century ago, Spearman (1904) discovered that when a battery of diverse cognitive tests is administered to a diverse group of people, there is consistent tendency for all the tests to be positively correlated with one another, producing what has been referred to as the “positive manifold”. Many studies since then have replicated this finding (Carroll, 1993; Jensen, 1998; Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004). Although the existence of general intelligence (*g*), in the sense of a statistical feature (a “positive manifold”), is a robust finding, it is less clear what the mechanisms are that support *g*. The best established candidate processes, as mechanistic substrates of *g*, are processing speed (Deary, 2001) and working memory (Conway, Jarrold, Kane, Miyake, & Towse, 2007). It remains

important to discover other processes that might similarly contribute to intelligence.

The current study investigated associative learning as a potential additional candidate process, which might contribute to *g* over and above processing speed and working memory. Only recently has associative learning become a serious contender as a substrate of *g* (Alexander & Smales, 1997; Tamez, Myerson, & Hale, 2008; Williams, Myerson, & Hale, 2008; Williams & Pearlberg, 2006). There is good reason, however, to suspect that the ability to learn associations might support *g*. Intelligent behavior seems certain to require memory for patterns of associations among stimuli, and one of the original purposes of intelligence tests was to assess students' ability to learn (Binet & Simon, 1916). Relations among associative learning, general cognitive ability, and cognitive mechanisms that subservise general cognitive ability are thus of interest for both theoretical and historical reasons.

In the present study, we are considering associative learning as the ability to remember and voluntarily recall

* Corresponding author. Department of Psychology, Yale University, Box 208205, New Haven, CT, 06520-8205, USA.

E-mail address: scott.kaufman@yale.edu (S.B. Kaufman).

specific associations between stimuli. Although early studies found a weak or no relation between associative learning and general cognitive ability (Malmi, Underwood, & Carroll, 1979; Underwood, Boruch, & Malmi, 1978; Woodrow, 1938, 1946), the failure to find a relation seems likely to be due to the fact that the associative learning tests that were used in these studies were easy and thus unlikely to be related to complex cognition (Estes, 1970). Consistent with this hypothesis, a more difficult associative learning task, in which subjects were required to learn multiple response–outcome contingencies for each trial, appears to be more strongly associated with *g* than a simpler associative learning task involving associations between pairs of stimuli (Williams & Pearlberg, 2006).

To be confident that associative learning is indeed a substrate of *g*, it is important to demonstrate that individual differences in associative learning make a contribution to the prediction of *g* that is statistically independent of the contributions of other candidate mechanisms. Otherwise, it might be the case that associative learning showed zero-order correlations with *g* merely because of its relation to some other important mechanism, such as working memory or processing speed. To the extent that *g* relies on multiple separable processes that are at least partially independent of one another, each should provide some incremental contribution to *g*. By examining the incremental validity of elementary cognitive tasks (ECTs, Jensen, 1998) that tap candidate cognitive mechanisms, one can effectively address the question of whether associative learning provides incremental prediction of *g* above and beyond two of the most well studied ECTs, working memory and processing speed tests.

Working memory is the ability to maintain, update, and manipulate information in an active state, over short delays (in the range of seconds rather than minutes). Individuals differ in their working memory, and those with higher working memory are better able to control their attention so as to maintain their task goals in the presence of interference (Conway, Cowan, & Bunting, 2001; Kane, Bleckley, Conway, & Engle, 2001; Unsworth, Schrock, & Engle, 2004). Working memory is strongly correlated with *g* (Conway et al., 2007; Engle & Kane, 2004; Heitz, Unsworth, & Engle, 2004). There is convincing evidence for a mechanistic link between working memory and *g*: tasks assessing *g* and working memory engage shared neural substrates, in lateral prefrontal cortex (PFC) as well as left and right parietal regions (Gray, Chabris, & Braver, 2003; Gray & Thompson, 2004). At least one additional cognitive mechanism has been identified that is very strongly related to *g*, namely processing speed.

Processing speed involves the speed at which even simple operations can be performed. Higher-IQ subjects respond faster in simple and choice reaction time paradigms (Deary, Der, & Ford, 2001) and are faster at perceiving a difference between two similar line segments in experiments on inspection time (Deary, 2000; Grudnik & Kranzler, 2001). In the Horn–Cattell theory of intelligence (Horn & Cattell, 1966), processing speed was described as “perceptual speed” (*G*s), and, in Carroll’s three-stratum theory of intelligence, as “general speediness” (Carroll, 1993). Finally, analysis of the factor structure of subtests from the WAIS (the standard IQ test) has demonstrated that processing speed is one of four second level factors below *g* (Deary, 2001).

The strong link between processing speed and *g* has led some researchers to argue that differences in *g* are primarily a result of differences in overall efficiency and speed of the nervous system (Anderson, 1992; Jensen, 1998). Others have criticized this view, on the grounds that performance on tests of processing speed may be a function of vigilance or ability to avoid distraction, rather than mere neural efficiency (Mackintosh, 1998). In any case, it seems unlikely that processing speed is the central mechanism underlying intelligence because measures of processing speed (*G*s) tend to load less strongly on *g* than other cognitive tests (Deary, 2001).

The possibility remains open that working memory and processing speed make separable, statistically independent contributions to *g*. Processing speed accounts for the link between working memory and *g* in some studies (Fry & Hale, 1996; Jensen, 1998; Kail & Salthouse, 1994; Salthouse, 1996), while others have found that working memory is the primary predictor of *g*, even while controlling for processing speed (Carpenter, Just, & Shell, 1990; Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen, 1996; Kyllonen & Christal, 1990). Conway et al. (2002) argue that these conflicting conclusions result from the use of processing speed tasks with different levels of working memory demand. At any rate, it is generally agreed that working memory is not identical to *g* (Ackerman, Beier, & Boyle, 2005; Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & Süß, 2005), leaving room for other processes to contribute additionally to *g*.

Associative learning and working memory correlate at the behavioural level of analysis (DeYoung, Peterson, & Higgins, 2005), and both appear to engage the PFC. However, working memory typically recruits dorsolateral areas of the PFC in Brodmann areas 9 and 46 (Petrides, 1995, 2000), whereas associative learning engages adjacent, more posterior frontal regions, in Brodmann areas 6 and 8 (Petrides, Alivisatos, Evans, & Meyer, 1993). The fact that the neural correlates of associative learning and working memory appear to be separable suggests that the two processes might make distinct contributions to intelligence. In support of this idea, a recent study found that learning of three-term contingencies predicted performance on Ravens Advanced Progressive Matrices (RAPM), a good measure of *g*, above and beyond working memory (Tamez et al., 2008). Further, once the variance between learning and *g* was accounted for, working memory no longer made a unique contribution to *g*. A limitation of this study, as well as that of Williams & Pearlberg (2006), is the analysis of single observed measures of *g* and learning rather than latent variables that model the shared variance of relevant tests and exclude error and unique method variance. Williams and Pearlberg (2006) found a more complex measure of learning to predict *g* more strongly than a simpler one, but the variance shared by both tasks may provide a better assessment of learning than either task alone. The current study was designed to overcome this limitation.

1.1. The present study

We administered the two associative learning tasks used by Williams and Pearlberg, (2006). Having examined these tasks individually, we proceeded to investigate associative learning as a latent construct by modeling their shared

variance. We used structural equation modeling to test the hypothesis that associative learning would predict *g* while controlling for working memory and processing speed. Support for this hypothesis would demonstrate that associative learning is a promising candidate as an elementary cognitive process contributing to *g*, separable from two of the best-established ECTs, working memory and processing speed.

2. Method

2.1. Participants

The 169 participants (54 males and 115 females) included in the analysis were aged 16–18 years, and attended a selective Sixth Form College (which takes high-achieving students who are in their last 2 years of secondary education) in Cambridge, England. Data were collected for 12 more participants, but 2 were removed from the analysis because their RAPM scores were below chance, 1 participant was removed due to obvious lack of effort (frequent chatting), and 9 other participants were removed because they failed to complete all tasks that were markers for one or more of the latent variables.

2.2. Procedure

Tests were administered in groups at PC desktop terminals during the course of three 1.5-h sessions. Whenever possible, all participants received all tests in the same order. Each participant earned £20 for their participation in all three testing sessions.

2.3. Associative learning tasks

2.3.1. Three-term contingency learning (Williams & Pearlberg, 2006)

The Three-Term Contingency Learning (3-Term) task consists of four learning blocks, each followed immediately by a test block. In each learning block, participants were presented with 10 unique words. Each word was associated with three different words, contingent on a key press. The participants' task was to learn the word associated with each stimulus–response pair. For instance, on one trial the word “LAB” might show on the screen with the letters “A”, “B”, and “C” listed underneath. When participants selected “A”, they saw one association (e.g., PUN), when they selected “B”, they saw a second association (e.g., TRY), and when they selected “C” they saw a third association (e.g., EGG). The duration of exposure to each association was self-paced (max 2.5 s) with changeover intervals set at 0.2 s. After the single presentation of all ten stimulus words with the 30 outcome words, subjects were immediately presented with a test block.

The test blocks were identical to the learning blocks with one exception: instead of typing the letters “A”, “B”, or “C” to produce the outcome words on the screen, a stimulus word appeared on the screen along with one of “A”, “B”, or “C”, and participants were required to type in the outcome word corresponding to that stimulus–response pair. Together with feedback on their answer, the correct association was shown to the participants until they pressed “ENTER”, when the next

stimulus word was presented. Once the test block was completed, participants immediately moved to a second learning block in which the same stimulus words were presented in a different order. Across the four test blocks, possible overall scores ranged from 0 to 120.

2.3.2. Paired–associates (PA) learning (Williams & Pearlberg, 2006)

In this task, participants were presented with 30 pairs of words. A cue word was presented until the participant pressed ENTER, or until 2.5 s elapsed, after which the cue's pair appeared on the screen. They then remained together on screen, again until the participant pressed ENTER, or until 2.5 s elapsed, after which both disappeared and the next cue word was displayed. The test phase was identical to training, except instead of pressing “ENTER” to view the second word of each pair, subjects were required to type that word. Together with feedback on their answer, the correct association was shown to the participant until they pressed “ENTER”, when the next word cue was presented. Once the test phase was completed, participants immediately moved to a second learning block in which the same stimulus words were presented in a different order. In total, there were four learning and four test blocks, with possible overall scores ranging from 0 to 120.

2.4. General cognitive ability tests

To create a good latent *g* factor we used one verbal test, one perceptual reasoning test, and one mental rotation test. Using one of the largest batteries of cognitive tests ever collected, Johnson and Bouchard (2005) demonstrated that, below the *g* factor, there are three separable second-stratum domains of cognitive ability: verbal, perceptual, and mental rotation. Use of one test from each domain should produce a well balanced *g*.

2.4.1. Raven's advanced progressive matrices test, set II (RAPM)

The RAPM (Raven, Raven, & Court, 1998) is a measure of abstract perceptual reasoning. Each item consists of a 3×3 matrix of geometric patterns with the bottom right pattern missing. The participants' task is to select the option that correctly completes the matrix. There are eight alternative answers for each item. The test is presented in increasing order of difficulty. After two practice items with feedback, participants were then given 45 min to complete 36 items.

2.4.2. DAT verbal reasoning test

The verbal reasoning section of the Differential Aptitudes Test (DAT-V, The Psychological Corporation, 1995) was administered to each participant. Each problem consisted of a sentence with two words missing, and participants chose a pair of words from the answer options that were related to the words in the sentence in some way. After two practice items, participants had 15 min to complete 40 problems.

2.4.3. Mental rotations test, set A (MRT-A)

The MRT-A (Vandenberg & Kruse, 1978) contains 24 problems and measures mental rotation ability, which appears to be a distinct component of intelligence at the same level as verbal ability and perceptual ability (Johnson &

Bouchard, 2005). Each problem in the MRT-A shows a three-dimensional target figure paired with four choice figures, two of which are rotated versions of the target figure. To score a point, both rotated versions must be identified. After two practice items with feedback and an explanation, the first 12 problems were attempted in 4 min with a 2 min break before attempting the second 12 in another 4 min. The maximum score is 24.

Mean scores on the three cognitive ability measures (RAPM, DAT-V, and MRT-A) suggested a mean IQ for the entire sample in the range of 100 to 110.

2.5. Processing speed tests

2.5.1. Verbal speed test (*Speed-V*)

An English adaptation of a sub-test from the Berlin model of Intelligence Structure (BIS; Jaeger, 1982, 1984). The task was to fill in the missing letter from a 7-letter word; 60 s were given to complete the 57 items. The score is the number completed correctly in 60 s.

2.5.2. Numerical speed test (*Speed-N*)

The Speed of Information Processing sub-test from the British Ability Scales (Elliot, 1996). The task was to cross out the highest number in each row of five numbers; 60 s were given to complete 48 items. The score is the number completed correctly in 60 s.

2.5.3. Figural speed test (*Speed-F*)

Digit-Symbol, Coding, a sub-test of the WAIS-R that loads on the “processing speed” factor (Deary, 2001). The test was to enter the appropriate symbol (given by a key at the top of the form) beneath a random series of digits; 90 s were given to complete 93 items. The score is the number completed correctly in 90 s.

2.6. Working memory

2.6.1. Operation span task (Turner & Engle, 1989)

The Operation Span (Ospan) task requires participants to store a series of unrelated words in memory while simultaneously solving a series of simple math operations, such as “Is $(9/3) - 1 = 1$?”. After participants selected the answer, they were presented with a word (e.g., DOG) to recall. Then participants moved on to the next operation–word string. This procedure was repeated until the end of a set, which varied from two to six items in length. Participants were then prompted to recall all the words from the past set in the same order in which they were presented by typing each word into a box, and using the up and down arrow keys on the keyboard to cycle through the boxes.

Before the main task, participants encountered three practice problems with set size two, where they received feedback about their performance. During these practice trials, we calculated for each participant how long it took them to solve the math operations. Consistent with the methodology of the Automated Ospan task (Unsworth, Heitz, Schrock, & Engle, 2005), we did this to control for individual differences in the time required to solve the math operations. Their mean performance time to solve the equations, plus

2.5 SD was used as the time limit for the presentation of the math equations during the main task.

The Ospan score is the sum of all correctly recalled words in their correct positions. The number of operation word-pairs in a set was varied between two, three, four, five, and six with three sets of each. Overall score could range from 0 to 60. Prior research has demonstrated significant correlations between Operation Span and *g* (e.g., Unsworth & Engle, 2005a) and a high loading of Operation Span on a general working memory factor (Kane et al., 2004).

2.7. Missing values

Some participants were missing values for certain variables, which we estimated using expectation–maximization based on the other markers of the relevant latent construct. Due to computer error, values were missing for 13 participants for one of the three markers of *g*. Data from the other two markers of *g* were used to impute 11 missing RAPM values, 1 missing DAT-V value, and 1 MRT-A value. For Speed-F, 10 participants did not follow the directions correctly and their scores could not be included in the analysis. Therefore, we used data from the other two markers of processing speed (Speed-V and Speed-N) to impute 10 missing values on Speed-F. Finally, due to a computer error, performance on the last trial of PA was not recorded for one participant. Since this participant achieved a maximum score on the third trial, we estimated that performance on the last trial was also a perfect score.

3. Results

3.1. Psychometric properties of the associative learning tasks

Table 1 shows the descriptive statistics for each trial of learning on both PA and 3-Term. The first block of learning on 3-Term had both an exceptionally low mean of 3.19 out of 30 correct and a standard deviation about half that of the other three learning trials. Similarly, the first block of learning on PA displayed a mean of 11.85 out of 30 correct, also low compared to performance on the other three blocks of learning on PA. Even so, performance on the first block of PA is significantly higher than performance on the first block of 3-Term [$t(168) = -19.11, p < .001$], suggesting a faster acquisition function and lower difficulty of the PA task relative to 3-Term.

To investigate the relation between *g* and each block of learning, we calculated each participant’s *g* score by assessing

Table 1
Descriptive statistics for learning trials on 3-Term and PA ($N = 169$).

	Mean	S.D.	Min	Max	Correlation with <i>g</i>
3-Term trial 1	3.28	3.50	0	21	.19*
3-Term trial 2	9.27	6.61	0	30	.34**
3-Term trial 3	14.62	7.91	0	30	.33**
3-Term trial 4	18.82	8.17	0	30	.33*
PA trial 1	11.85	6.65	1	28	.19*
PA trial 2	20.54	7.36	2	30	.29**
PA trial 3	24.08	6.73	2	30	.27**
PA trial 4	25.79	5.85	5	30	.29**

* $p < .05$; ** $p < .01$.

the common variance across RAPM, DAT-V, and MRT-A using Principal Axis Factoring. Performance on the four blocks of each learning task and the relation of g to performance across blocks was analyzed using a repeated measure GLM. For both the 3-Term and PA learning tasks, there was a significant main effect for block [3-Term: $F(3,165)=277.56$, $p<.001$; PA: $F(3,165)=370.45$, $p<.001$] and a significant $g \times$ block interaction [3-Term: $F(3,165)=7.21$, $p<.001$; PA: $F(3,165)=2.88$, $p<.05$]. This interaction indicates that for both learning tasks, g is more strongly associated with some learning blocks than others. Table 1 shows the correlation between g and each block of learning on both PA and 3-Term. The correlation of each block of learning with g stabilizes after the first block of learning, suggesting that the first block of learning for both measures of associative learning may be too difficult for reliable individual differences to emerge. For this reason, further analyses will exclude the first block on both 3-Term and PA.

3.2. Associative learning and g

Cronbach's alpha was used to estimate the reliability of the two associative learning tasks. For both measures of associative learning, reliability was calculated across blocks 2, 3, and 4. Both 3-Term ($\alpha=.93$) and PA ($\alpha=.95$) showed high reliability. The correlation matrix for scores on observed variables appears in Table 2, with descriptive statistics for each test. The correlation between the total scores on 3-Term and PA was significant and high, suggesting that the two measures were engaging the same ability, or at least similar processes. It is also noteworthy that 3-Term and PA were both significantly correlated with RAPM.

To test predictions about the independent prediction of g by associative learning (AL), working memory (WM), and processing speed (Gs), we used structural equation modeling. The model was analyzed using Amos 7.0 (Arbuckle, 2006) with maximum likelihood estimation. (Appendix A includes the full covariance matrix used to fit the model in Fig. 1).

The shared variance of blocks 2, 3, and 4 of the 3-Term test phases formed the latent variable "3-Term," and the shared variance of blocks 2, 3, and 4 of the PA test phases formed the latent variable "PA." These two latent variables then formed the latent AL variable.

The shared variance across Ospan trials of set size two, three, four, five, and six formed the latent variable representing WM. The shared variance across Speed-V, Speed-F, and Speed-N formed the latent variable representing Gs. The shared variance across RAPM, DAT-V, and MRT-A formed the latent variable representing g .

Prior to fitting the structural model predicting g (Fig. 1), we fit a model simply allowing the predictor variables to correlate with each other and with g , in order to assess their zero-order associations. The fit of this model was almost identical to that in Fig. 1. Correlations among the latent variables appear in Table 3. AL, WM, and Gs are significantly correlated with g . Although correlations are not shown in Fig. 1, all predictors of g were allowed to correlate with each other.

In the model shown in Fig. 1, AL, WM, and Gs all make significant independent contributions to g . The model accounts for 40% of the total variance in g . Also listed in Fig. 1 is the χ^2 test for significant discrepancies between the predicted and observed covariance matrices, as well as the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). A significant χ^2 does not necessarily indicate poor fit because the χ^2 value is sensitive to sample size (Kline, 2005). Other fit indices are designed to surmount this limitation. CFI and TLI values over .90 indicate adequate fit and values of .95 or higher indicate close fit (Kline, 2005). RMSEA values less than 0.08 indicate acceptable fit, while values of 0.05 or less indicate close fit (Kline, 2005). The $p_{(close)}$ statistic indicates whether the RMSEA value is significantly greater than 0.05. The fit indices reported in Fig. 1 reveal that the structural model provides a good fit to the data.

4. Discussion

The main aim of the study was to investigate the contribution of associative learning to g , above and beyond working memory and processing speed. In line with this goal, we first examined the psychometric properties of two associative learning tasks – three-term contingency learning and paired-associates learning. Analyses indicated that 3-Term was a more difficult task than PA overall, but both tasks demonstrated a relatively low score on the first block of

Table 2
Correlations, means, and standard deviations of observed variables ($N=169$).

Measure	1	2	3	4	5	6	7	8	9
1. RAPM (/36)	–								
2. DAT-V (/40)	.53**	–							
3. MRT-A (/24)	.59**	.43**	–						
4. Ospan (/24)	.30**	.42**	.24**	–					
5. Speed-V (/57)	.17**	.24**	.14	.22**	–				
6. Speed-F (/90)	.24**	.15*	.17*	.15	.24**	–			
7. Speed-N (/48)	.25**	.10	.21**	.10	.14	.51**	–		
8. 3-Term (/90)	.32**	.35**	.21**	.21**	.09	.16*	.02	–	
9. PA (/90)	.31**	.23**	.14	.14	.17*	.17*	-.05	.64**	–
Mean	21.72	24.52	13.21	44.53	41.36	64.70	30.90	42.70	70.42
S.D.	5.53	5.86	5.34	7.61	9.05	10.50	4.19	21.37	19.15
Reliability	.81	.78	.85	.73	.60	.60	.60	.93	.95

Notes: All reliability analyses are alpha coefficients, except for the three processing speed tests, in which the Spearman–Brown split-half coefficient was calculated across all three tests. The parenthetical next to each test refers to the total score for each test.

* $p<.05$, ** $p<.01$.

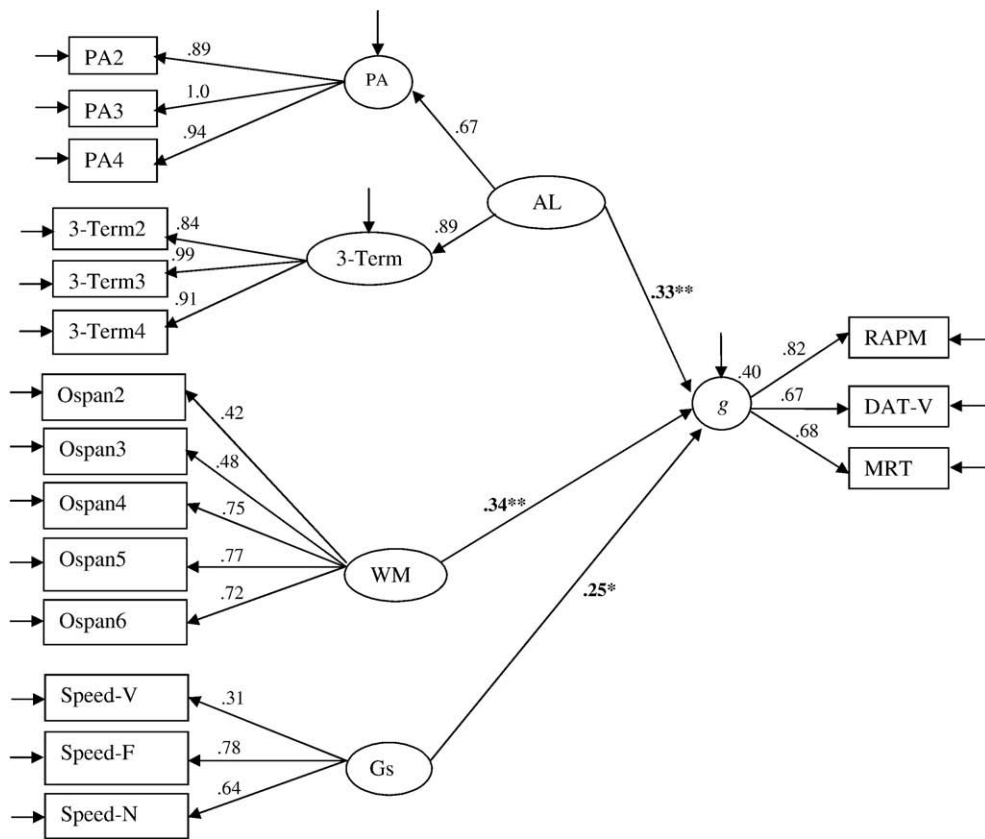


Fig. 1. Associative learning, working memory capacity (WM), and processing speed (Gs) independently predict g. See Table 3 for correlations among latent predictors. $N = 169$. $\chi^2 = 175.21$, $df = 111$, $p < .001$, CFI = .96, TLI = .95, RMSEA = .059, $p_{(close)} = .189$. * $p < .05$, ** $p < .01$.

learning. Further, while g was significantly associated with learning on each block, correlations with g changed significantly over the four blocks, becoming stronger on blocks 2, 3, and 4 for both 3-Term and PA. We therefore excluded the first block of learning from both measures of associative learning from our structural analysis. Any study that assesses the relation between associative learning and g must consider the psychometric properties of the tasks, to ensure that they are good measures of individual differences.

Structural equation modeling showed that associative learning, working memory, and processing speed all made statistically independent contributions to g . This finding suggests that each of these elementary cognitive processes may represent a mechanism that contributes differentially to general intelligence. The current study is consistent with

three recent studies that have demonstrated a link between associative learning and g . Firstly, Williams and Pearlberg (2006), found that 3-Term learning was related to RAPM but was not significantly related to various measures of processing speed, which matches our results (although their results differ from ours in that they did not find a significant association between PA and RAPM). Secondly, a more recent study conducted by Tamez, Myerson, and Hale (2008) replicated the correlation between 3-Term and RAPM, but also found that 3-Term correlated with RAPM even after controlling for working memory. Thirdly, Alexander and Smales (1997) found that the composite of various verbal and nonverbal learning tasks was correlated at .49 with the composite of various measures of general ability. The current study is consistent with these studies but goes further, in that it indicates the existence of a general associative learning ability factor that is related to a latent g factor, even after controlling for both processing speed and working memory.

The finding of a significant correlation between working memory and g is consistent with a growing and consistent literature on the strong relation between variation in working memory and general cognitive ability (Conway et al., 2007; Engle & Kane, 2004; Engle et al., 1999; Heitz et al., 2004; Kyllonen, 1996; Kyllonen & Christal, 1990). The finding of a significant relation between processing speed and g is also well supported by a large literature (Deary et al., 2001; Jensen, 2006).

Table 3
Correlation matrix of latent variables in structural model ($N = 169$).

Measure	1	2	3	4
1. g	–			
2. Working memory (WM)	.48**	–		
3. Processing speed (Gs)	.38**	.24*	–	
4. Associative learning (AL)	.45**	.22*	.17	–

* $p < .05$, ** $p < .01$.

It should be noted that even though WM significantly predicted *g* independently of the other variables, the effect size of association between WM and *g* in our SEM models is lower than what has been reported elsewhere (e.g., Conway et al., 2002; Kyllonen & Christal, 1990). This difference may be due to the fact that only one test of working memory was administered in the current study. Adding more varied indicators to the WM latent variable would most likely have increased the relation between WM and *g*. Nonetheless, the measure of working memory administered in the current study, Operation Span, displayed similar zero-order correlations with RAPM as in other studies (Conway et al., 2002; Engle et al., 1999; Kane et al., 2004; Unsworth & Engle, 2005b), including a study that found an association between the 3-Term learning task and RAPM, when controlling for WM (Tamez et al., 2008).

A parallel concern to the use of only one working memory task is the use of multiple associative learning tasks, creating a model that is slightly unbalanced in the number of markers for each of our three predictors, with the most markers for associative learning. To address this concern we tested two additional structural models, one excluding the PA task and one excluding the 3-Term task. In both of these models, with only one task used to create a latent associative learning variable, the overall pattern of findings remained the same as in Fig. 1, with all three latent variables significantly predicting *g*. In both of these models, the paths from WM and AL to *g* were not significantly different from each other.

Although more thorough measurement of WM might have reduced the amount of variance in *g* explained by AL, it is equally plausible that more thorough measurement of associative learning could have reduced the variance explained by WM. Indeed, Tamez, Myerson, and Hale (2008) found that Ospan's correlation with RAPM was no longer significant after controlling for 3-Term, suggesting that it is an open question whether WM or AL is the primary predictor of *g*. And of course there is a third alternative, suggested by the present study, which is that both WM and AL predict *g* independently. Future studies will hopefully provide a more thorough test of this hypothesis by including more measures of both WM and AL. Constructing additional associative learning tasks that tap into a general associative learning ability factor will be an important step in this direction (Williams et al., 2008).

The independent prediction of *g* by both working memory and processing speed is inconsistent with work by Conway et al. (2002), who found that processing speed no longer predicted *g* after controlling for working memory. A comparison of the processing speed tasks administered in their study and the current study shows that very similar tasks were administered, with one being identical (Digit–Symbol Coding). There are at least four possible reasons for the discrepancy between our findings and theirs. First, their assessment of *g* is not as comprehensive as ours, as they utilized only the RAPM and one other very similar test. Thus, their *g* is most closely related to the perceptual ability component of the second-stratum factors identified by Johnson and Bouchard (2005). Arguing against this explanation, however, is the fact that when we re-ran our structural model using observed RAPM scores as our criterion variable instead of *g*, all three latent predictors remained significant.

The second possible reason for the discrepancy is that they included a latent short-term memory variable as a predictor in addition to working memory and processing speed. Short-term memory was correlated with processing speed and could have suppressed the latter's association with *g*. The third possibility relates to our measurement of working memory, discussed above. If we had used additional measures of WM, WM might have related more strongly to *g* and this additional variance predicted by WM might have rendered that predicted by processing speed non-significant. Even if this were the case, however, it might simply indicate that processing speed is a lower-level mechanism that contributes to WM as well as to *g*. The fourth possible reason for the discrepancy is a difference in the developmental stage of the participants. Their sample consisted of college students, who were older than the current sample of Sixth Form students.

Before concluding, we should note that the associative learning ability assessed in this study is distinct from the associative paradigms that are used to assess implicit learning and tacit knowledge (e.g., Gebauer & Mackintosh, 2007; Reber, Walkenfeld, & Herndstadt, 1991). In implicit learning, associations between stimuli are not acquired voluntarily, but rely on mere exposure without awareness of the association. In the associative learning tasks employed here, by contrast, subjects consciously and voluntarily remember associations. Different mechanisms are likely to be involved in explicit versus implicit associative learning, and our findings should not be assumed to generalize to implicit learning.

5. Conclusion

The results of the current study add to a growing literature on the existence of multiple cognitive mechanisms that support general cognitive ability (Sternberg & Pretz, 2005). Our findings suggest that multiple cognitive processes – including the abilities to process information quickly, to maintain, update, and manipulate information in working memory, and to learn specific associations between stimuli – should contribute to performance on any highly *g*-loaded task. Identification of separable elementary cognitive mechanisms that support *g* should further attempt to develop neurobiological theories of intelligence. Such theories may help to resolve current debates regarding the nature of the mechanisms underlying *g* (e.g., Colom, Francis, Quiroga, Shih, & Flores-Mendoza, 2008). Evidence exists already that working memory and associative learning rely on different regions of the PFC (Petrides, 1995, 2000; Petrides et al., 1993), and processing speed seems likely to be determined by a distinct set of biological parameters that are not yet known. The investigation of the precise number and nature of the mechanisms that underlie *g* remains a promising line of research.

Acknowledgements

The authors would like to thank Sheila Bennett for her kind assistance in recruiting participants, Jim Blair and Nikhil Srivastava for computer support, and the administration at Hills Road Sixth Form College for the use of their facilities.

Appendix A. Full covariance matrix used to fit SEM model (N = 169).

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
RAPM	–																			
DATV	.53	–																		
MRTA	.59	.43	–																	
SPEEDV	.17	.24	.14	–																
SPEEDF	.24	.15	.17	.24	–															
SPEEDN	.25	.10	.21	.14	.51	–														
Ospan2	.12	.25	.11	.20	.11	.08	–													
Ospan3	.24	.21	.18	.19	.09	.03	.33	–												
Ospan4	.23	.41	.23	.19	.12	.13	.35	.38	–											
Ospan5	.27	.34	.21	.22	.16	.11	.27	.37	.56	–										
Ospan6	.21	.30	.16	.10	.08	.03	.25	.26	.55	.60	–									
ThreeTerm1	.17	.18	.12	.86	.23	.05	.07	.10	.08	.16	.17	–								
ThreeTerm2	.32	.30	.21	.84	.15	–.01	.07	.11	.04	.16	.15	.72	–							
ThreeTerm3	.30	.35	.20	.10	.14	.01	.12	.10	.07	.16	.20	.57	.84	–						
ThreeTerm4	.30	.34	.17	.07	.16	.05	.15	.14	.11	.20	.20	.46	.74	.90	–					
PA1	.20	.20	.07	.14	.18	–.04	.05	.08	.05	.09	.19	.48	.58	.67	.59	–				
PA2	.30	.25	.12	.16	.19	–.02	.05	.09	.03	.12	.17	.33	.52	.65	.65	.84	–			
PA3	.28	.20	.14	.16	.15	–.05	.03	.10	.01	.15	.12	.24	.45	.58	.64	.69	.89	–		
PA4	.30	.22	.15	.16	.14	–.08	.04	.11	.02	.18	.10	.21	.44	.55	.61	.60	.82	.94	–	

References

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs. *Psychological Bulletin*, *131*, 30–60.
- Alexander, J. R., & Smales, S. (1997). Intelligence, learning and long-term memory. *Personality and Individual Differences*, *23*, 815–825.
- Anderson, M. (1992). *Intelligence and development: A cognitive theory*. Malden, MA: Blackwell Publishing.
- Arbuckle, J. L. (2006). *Amos 7.0 (Build 1140): Amos Development Corporation*.
- Binet, A., & Simon, T. (1916). *The development of intelligence in children*. Baltimore, MD: Williams & Wilkins (E. S. Kite, Trans.).
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the raven progressive matrices test. *Psychological Review*, *97*, 404–431.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Colom, R., Francis, J. A., Quiroga, M.Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, *36*, 584–606.
- 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.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.
- Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (2007). *Variation in working memory*. New York, NY: Oxford University Press.
- Corporation, The Psychological (1995). *DAT for selection-verbal reasoning (European adaptation)*. Orlando, FL: Harcourt Brace and Company.
- Deary, I. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford, UK: Oxford University Press.
- Deary, I. (2001). Human intelligence differences: A recent history. *Trends in Cognitive Sciences*, *5*, 127–130.
- Deary, I., Der, G., & Ford, G. (2001). Reaction times and intelligence differences: A population-based cohort study. *Intelligence*, *29*, 389–399.
- DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*, 825–858.
- Elliot, C. D. (1996). *British Ability Scales II*. Windsor, Berkshire, UK: NFER-Nelson.
- 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: Advances in research and theory*, Vol. 44. (pp. 145–199) New York, NY: Elsevier.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.
- Estes, W. K. (1970). *Learning theory and mental development*. New York, NY: Academic Press.
- Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237–241.
- Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 34–54.
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.
- Gray, J. R., & Thompson, P. M. (2004). Neurobiology of intelligence: Science and ethics. *Nature Reviews Neuroscience*, *5*, 471–482.
- Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence*, *29*, 523–535.
- Heitz, R. P., Unsworth, N., & Engle, R. W. (2004). Working memory capacity, attention control, and fluid intelligence. In I. O. Wilhelm, & R. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 61–77). New York, NY: Sage.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*, 253–270.
- Jaeger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. [Multimodal classification of cognitive abilities: An experimentally controlled development of a descriptive model of intelligence structure]. *Diagnostica*, *28*, 195–225.
- Jaeger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven [Research in the structure of intelligence: Competing models, new developments, perspectives]. *Psychologische Rundschau*, *35*, 21–35.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. St. Louis, MO: Elsevier Science.
- Johnson, W., & Bouchard, T. J., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.
- Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, *32*, 95–107.
- Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199–225.
- 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*, 169–183.
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 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, 189–217.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.
- Kyllonen, P. C. (1996). Is working memory capacity Spearman's *g*? In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 49–75). Mahwah, NJ: Erlbaum.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14, 389–433.
- Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford, UK: Oxford University Press.
- Malmi, R. A., Underwood, B. J., & Carroll, J. B. (1979). The interrelationships among some associative learning tasks. *Bulletin of the Psychonomic Society*, 13, 121–123.
- Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 61–65.
- Petrides, M. (1995). Impairments on non-spatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience*, 15, 359–375.
- Petrides, M. (2000). Dissociable roles of mid-dorsolateral prefrontal and anterior inferotemporal cortex in visual working memory. *The Journal of Neuroscience*, 20, 7496–7503.
- Petrides, M., Alivisatos, B., Evans, A. C., & Meyer, E. (1993). Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proceedings of the National Academy of Sciences of the U.S.A.*, 90, 873–877.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for raven's progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.
- Reber, A. S., Walkenfled, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 888–896.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, 15, 201–293.
- Sternberg, R. J., & Pretz, J. E. (2005). *Cognition and intelligence: Identifying the mechanisms of the mind*. New York, NY: Cambridge University Press.
- Tamez, E., Myerson, J., & Hale, S. (2008). Learning, working memory, and intelligence revisited. *Behavioural Processes*, 78, 240–245.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127–154.
- Underwood, B. J., Boruch, R. F., & Malmi, R. A. (1978). Composition of episodic memory. *Journal of Experimental Psychology: General*, 107, 393–419.
- Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, 33, 213–220.
- Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between operation span and raven. *Intelligence*, 33, 67–81.
- Unsworth, N., Heitz, R., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 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, & Cognition*, 30, 1302–1321.
- Vandenberg, S. G., & Kruse, A. R. (1978). Mental rotations: Group tests of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599–604.
- Williams, B. A., Myerson, J., & Hale, S. (2008). Individual differences, intelligence, and behavior analysis. *Journal of the Experimental Analysis of Behavior*, 90, 219–231.
- Williams, B. A., & Pearlberg, S. L. (2006). Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence*, 34, 177–191.
- Woodrow, H. (1938). The relation between abilities and improvement with practice. *Journal of Educational Psychology*, 29, 215–230.
- Woodrow, H. (1946). The ability to learn. *Psychological Review*, 53, 147–158.