Mem Cogn
DOI 10.3758/s13421-014-0473-3

Why is working memory capacity related to matrix reasoning tasks?

Tyler L. Harrison · Zach Shipstead · Randall W. Engle

Abstract One of the reasons why working memory capacity is so widely researched is its substantial relationship with fluid intelligence. Although this relationship has been found in numerous studies, researchers have been unable to provide a conclusive answer as to why the two constructs are related. In a recent study, researchers examined which attributes of Raven's Progressive Matrices were most strongly linked with working memory capacity (Wiley, Jarosz, Cushen, & Colflesh, Journal of Experimental Psychology: Learning, Memory, and Cognition, 37, 256–263, 2011). In that study, Raven's problems that required a novel combination of rules to solve were more strongly correlated with working memory capacity than were problems that did not. In the present study, we wanted to conceptually replicate the Wiley et al. results while controlling for a few potential confounds. Thus, we experimentally manipulated whether a problem required a novel combination of rules and found that repeated-rule-combination problems were more strongly related to working memory capacity than were novel-rule-combination problems. The relationship to other measures of fluid intelligence did not change based on whether the problem required a novel rule combination.

Keywords Working memory · Intelligence

Working memory consists of a system of temporary memory stores, the functions of retrieval and maintenance into and out of those stores, and the executive attention necessary to the performance of these functions. Working memory *capacity* (WMC) refers to the effectiveness of the working memory system for a given individual. One of the most ubiquitous and important findings in the study of WMC is its strong relationship to fluid intelligence (Gf), the ability to solve novel reasoning problems. The nature of this relationship has been heavily debated, with some researchers arguing that WMC and Gf are essentially the same construct (e.g., Martínez et al., 2011), and others claiming that the two constructs are clearly separable (e.g., Ackerman, Beier, & Boyle, 2005; Heitz et al., 2006; Kane, Hambrick, & Conway, 2005). Although many studies have demonstrated the relationship between WMC and Gf (e.g., Engle & Kane, 2004; Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004), the underlying cause of the correlation has not been identified. Researchers have suggested a number of alternative possibilities. Engle and Kane argued that executive attention is important to both reasoning ability and WMC. Others have suggested that simple memory maintenance can completely account for the relationship (Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Colom, Rebollo, Abad, & Shih, 2006).

These theories for the link between Gf and WMC have primarily been tested at the latent-construct level. Although this approach has many advantages (see Engle et al., 1999), researchers may be unable to analytically determine the reasons for the latent variable correlation. To fully understand why the latent relationship between WMC and Gf exists, a researcher may conduct both differential and experimental research to examine which attributes of tasks that measure Gf task can be manipulated in order to change the task's correlation to WMC. The attributes that lead to the strongest changes to the correlation with WMC would inform researchers about the cause of the WMC/Gf relationship. A few studies of this nature have been conducted with one of the most common measures of Gf, Raven's Advanced Progressive Matrices (RAPM; Raven, Raven, & Court, 1998). This measure is a matrix reasoning test in which

T. L. Harrison (✉) · R. W. Engle
School of Psychology, Georgia Institute of Technology,
654 Cherry Street, Atlanta, GA 30332, USA
e-mail: tharrison9@gatech.edu

Z. Shipstead
Arizona State University, Phoenix, AZ, USA

Published online: 21 October 2014

participants are given a 3 ×3 matrix of figures with the bottom-right figure missing (see Fig. 1 for an example problem). The figures form a pattern, and subjects must select the correct figure to fill in the missing spot with one of eight provided answer choices.

Raven's problems require varying numbers of rules to solve. For instance, one Raven's problem may require the subject to add two figures together to correctly solve the problem, whereas another may require subjects to progressively rotate a shape. Simple problems may require only one rule to solve, but more complex problems may require multiple rules. Carpenter, Just, and Shell (1990) suggested that WMC is positively correlated with Raven's because WMC is needed to maintain rules in mind to successfully answer Raven's problems. They argued that the importance of WMC for Raven's performance increases with the number of rules required for solving the problem. To test their claim, these researchers developed two computer simulations that only differed in how many rules they could hold in working memory. The computer model with higher working memory solved Raven's problems more accurately than did the computer model with lower working memory. The idea behind these two models was that if the models performed like two different groups of human subjects, then the difference between the two models might reflect the difference between the two groups of humans.

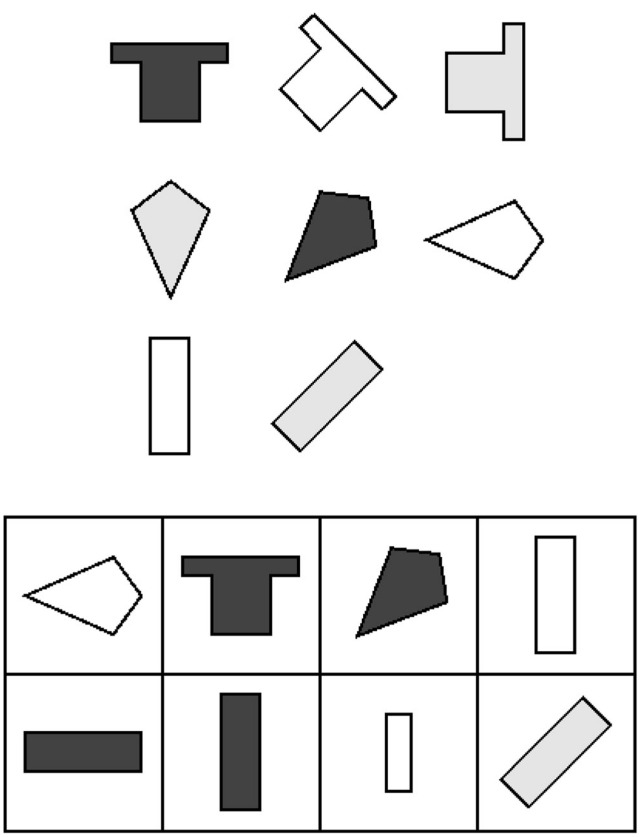

Fig. 1 Example of a matrix reasoning item similar to Raven's problems, created using the Sandia Generated Matrix Tool (Matzen et al., 2010)

Unsworth and Engle (2005) directly tested the claims of Carpenter et al. (1990) by examining the relationship between WMC and each individual problem of Raven's using human subjects. They found that WMC correlated just as highly to Raven's problems that required one rule to solve as to those that required multiple rules. Instead of the number-of-rules account, they suggested that WMC may be correlated with Raven's due to individual differences in proactive inference. Complex span tasks (i.e., empirically validated measures of WMC) require subjects to remember items for recall only from the current trial and to ignore items from previous trials. With each new trial, subjects must ignore items from previous trials. Studies have shown that the ability to manage proactive interference is important both to complex span performance and to the correlation between individual differences in WMC and Gf (Bunting, 2006; May, Hasher, & Kane, 1999). It is therefore reasonable to assume that proactive inference plays a role in solving Raven's problems. The rules of previously completed problems may interfere with subjects uncovering the rules of future problems.

Recently, Wiley, Jarosz, Cushen, and Colflesh (2011) proposed an *interference/distraction* account of the relationship between measures of WMC and Raven's problems (see also Jarosz & Wiley, 2012). They argued that WMC is useful in keeping previously learned rule combinations from interfering when solving Raven's problems that require a new rule combination. Therefore, they predicted that WMC should be highly correlated with Raven's problems that require a novel rule combination, and less correlated with problems that repeat the rule combination from a previous trial. In Wiley et al.'s first study, subjects completed both operation span (OSpan; Turner & Engle, 1989) and RAPM. Wiley et al. found that operation span correlated more highly with problems that required a novel rule combination (r = .39) than with repeated-rule problems (r = .26). Both the total of novel-rule-combination problems and subjects' overall RAPM scores were entered into a regression model to predict operation span performance. In this model, novel-rule problems were predictive of operation span performance, and the overall RAPM score was not. When the same model was used for the repeated-rule problems, only the overall RAPM score was predictive of operation span performance. From these data, Wiley et al. argued that the novel-rule-combination problems were the driving force behind the correlation between WMC and Raven's problems.

As further evidence of their hypothesis, Wiley et al. (2011) matched two different subsets of Raven's problems on difficulty. One set consisted of problems that could each be solved by a unique rule combination, and one set required only five different rule combinations to correctly solve all of the problems. In all, 25 subjects completed the novel-rule subset of

RAPM, and another 25 completed the common-rule subset. The novel-rule subset of RAPM strongly correlated with WMC (defined by a composite of the performance on operation span and reading span; $r = .62$), and the repeated-rules subset of RAPM did not correlate with WMC ($r = .02$). This substantial difference in these correlations provides further support for the *interference/distraction* account.

However, there are some limitations to Wiley et al.'s (2011) interpretations. For Study 1, RAPM was always given in the same order. The novel-rule-combination problems thus might relate more highly to WMC, not because these problems require a novel rule combination, but because of some idiosyncratic differences in these problems unrelated to the rule combinations. For Study 2, Wiley et al. had only 50 subjects, and due to the between-subjects design, roughly 25 subjects were given each subset of Raven's problems. One could then make the argument that the findings of Study 2 are the result of random fluctuations due to their small sample size. This is a concern that is particularly problematic when conducting correlational research (Schönbrodt & Perugini, 2013).[1]

There are also theoretical reasons why novel-rule problems should be *less* correlated with WMC than repeated-rule problems. For instance, Unsworth and Engle (2007) argued that WMC is helpful in retrieving items from secondary memory. From this perspective, having high WMC might allow subjects to remember more previous Raven's solutions. Therefore, when subjects receive a Raven's problem that requires a set of rules similar to the one on a previous problem, those with higher WMC could recall the solution to the previous problem to help solve the current problem. If this is the case, the correlation to WMC should be stronger for repeated-rule than for novel-rule problems. Wiley and colleagues have called this theory the *learning efficiency* account.

Additionally, it is unclear how the attribute of novel rule use affects the correlation of Raven's to other measures of Gf. Since Gf is thought to be the ability to uncover new relations in cognitive problems, it might be the case that novel-rule Raven's problems would be more highly related to other Gf measures than would repeated-rule Raven's problems. If this hypothesis were true, the results would be exciting from both theoretical and practical perspectives. Theoretically, we would know that Gf is an ability that is particularly related to solving novel problems. Practically, psychometricians might be able to develop tests that are better measures of Gf and, thus, better indicators of academic and job performance.

[1] We recently tried a within-subjects replication of Wiley et al.'s (2011) Study 2. With 99 subjects, we did not find that our novel-rule Raven's subset correlated to WMC more highly than did our repeated-rule subset. For more details of this study, go to http://englelab.gatech.edu/.

The present study

One potential confound in the Wiley et al. (2011) studies is that specific problems were always presented in the same order. It is possible that the novel-rule-combination problems in both of Wiley et al.'s studies correlate more highly with WMC because of idiosyncrasies of these specific problems that are unrelated to whether the problem requires the use of a novel rule combination. To eliminate this potential confound, we presented subjects with five pairs of problems, with each pair requiring the same rule combination to solve. We presented each pair of problems sequentially and manipulated the order of the problems so that, in one counterbalanced condition, a particular problem required a novel rule combination in order to solve it correctly, and in the other condition, the same problem required a repeated rule combination.

Method

Subjects

The subjects were 228 Georgia Tech students and members of the Atlanta community. These subjects completed two 2-h sessions and received a $30 check for each session they completed. Twenty of the subjects were eliminated from the data analysis because they did not complete their second session, leaving 208 subjects with complete data. The subjects completed a battery of cognitive tasks during each session that were part of a general screening procedure. We report only the tasks relevant to the present study. For a list of all of the tasks and the order in which they were presented, see Shipstead, Lindsey, Marshall, and Engle (2014, Table 1).

Table 1 Descriptive statistics

Task	Mean	Variance	Skewness	Kurtosis
OSpan	55.55	199.40	−0.98	0.70
SymSpan	26.06	74.64	−0.48	−0.23
RunSpan	39.04	152.11	0.13	−0.17
NovProblems	0.39	0.06	0.25	−0.45
RepProblems	0.43	0.06	0.26	−0.38
Letter Sets	14.94	20.20	−0.20	−0.39
NumSeries	8.60	9.61	−0.27	−0.39

OSpan = Operation Span; SymSpan = Symmetry Span; RunSpan = Running Span; NovProblems = Novel-Rule-Combination Problems; RepProblems = Repeated-Rule-Combination Problems; NumSeries = Number Series

Procedure

Operation span (OSpan; Unsworth, Heitz, Schrock, & Engle, 2005) The automated version of the operation span task was used for the present study (see Fig. 2). Subjects made judgments about whether or not an equation was correct and then were presented with a to-be-remembered letter. After three to seven mathematical judgments/letters, subjects attempted to recall the letters in the order in which they had been presented. The number of letters recalled in correct serial position was the dependent variable.

Symmetry span (SymSpan) In the automated version of the symmetry span task, subjects made a vertical symmetry judgment about a black-and-white figure and then were presented with a filled cell in a 4 ×4 matrix (see Fig. 2). After two to five matrix elements, subjects were required to recall the locations of the filled cells in the order they had been presented by clicking a mouse key in the appropriate cell. The number of matrix locations recalled in the correct order was the individual's score.

Running span (RunSpan; Broadway & Engle, 2010) Subjects were presented with a brief series of letters from three to nine in length. After the letters had been presented, subjects were required to recall a certain number of the most recent letters. For example, a subject might be cued to "recall the last three letters of the next set," and then presented with the letter set "QTJKD." The subject would have to enter the response "JKD." Subjects did not know how many letters would be presented for each trial. The number of letters recalled in the correct order was the score used for analysis.

Letter sets (Ekstrom, French, Harman, & Dermen, 1976) Subjects were presented with five sets of letters, with each set containing four letters. Subjects were instructed to find the rule that applied to four of the five letter sets, and then to indicate the letter set that violated the rule. There were 30 problems, and subjects had 5 min to complete them.

Number series (NumSeries; Thurstone, 1938) For this task, subjects were presented with a series of numbers and were instructed to identify the answer choice that was the next logical number in the sequence. There were 15 problems, and subjects had 4.5 min to complete them.

Raven's mixed Raven's mixed was a task designed to answer the question of whether matrix reasoning problems that require novel rule combinations correlate more strongly to WMC and Gf than do problems that require repeated rule combinations. For this task, subjects first completed two practice problems from Set I of RAPM. The instructions explained the solutions for both of the problems. After the practice problems, subjects were given 20 min to complete 14 matrix reasoning problems. All of the subjects completed the first four problems in the same order. These four problems (Problems 2, 4, 6, and 10 from RAPM) were given in order to familiarize subjects with matrix reasoning problems.

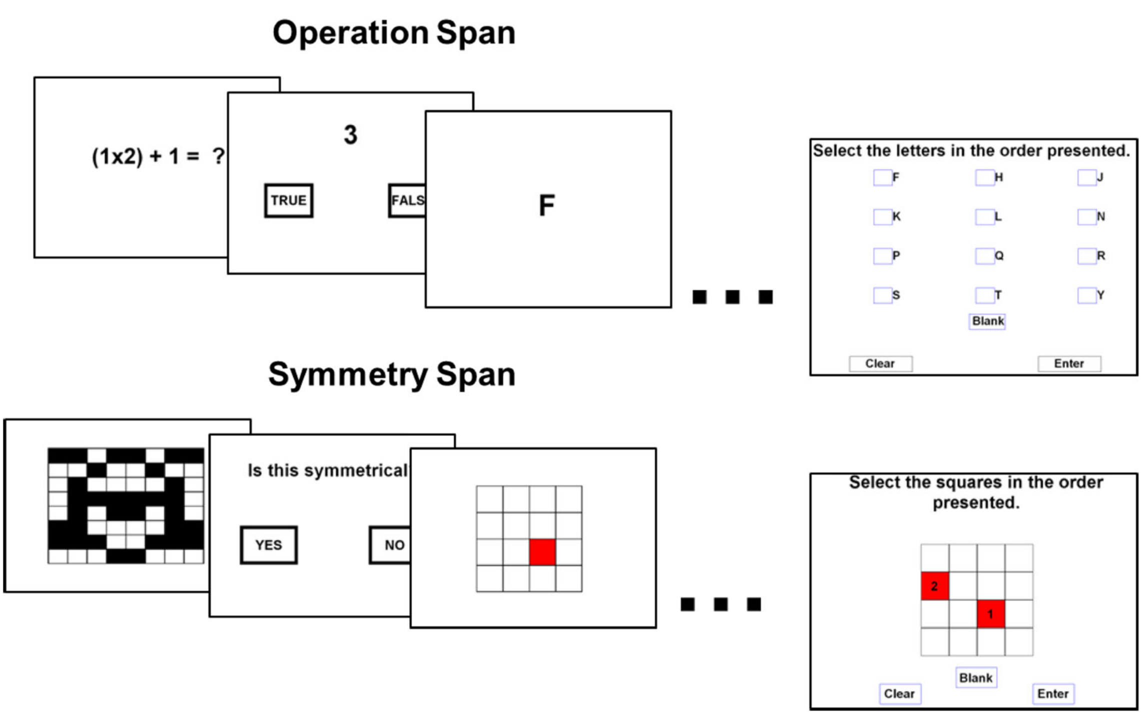

Fig. 2 Examples of the complex span tasks (adapted from Harrison et al., 2013)

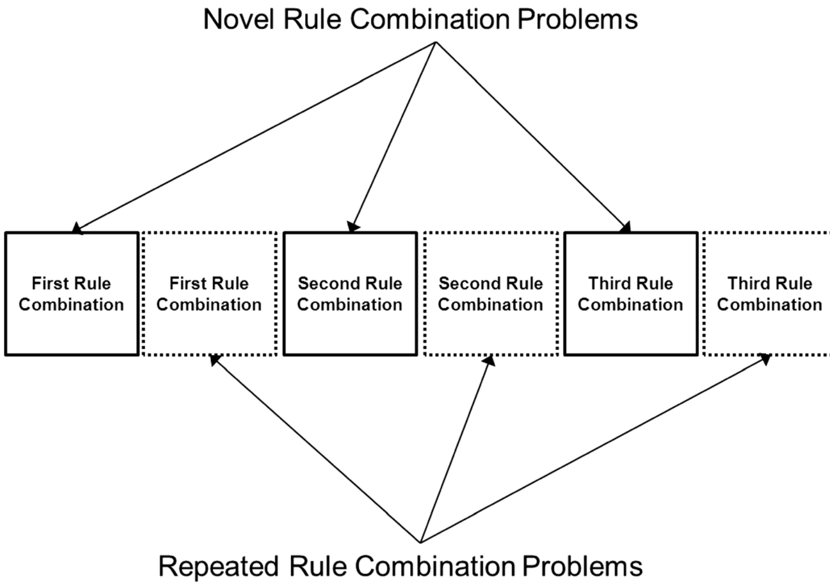

Fig. 3 Graphical representation of six problems from the Raven's mixed task. Each square represents a problem, and within each square the particular rule combination required is listed. In the counterbalanced condition, the locations for the solid-line problems and the dotted-line problems would be switched

Critically, these four problems did not require the same rule combinations as any of the last ten critical problems.

For the last ten problems, we selected five pairs of problems from either RAPM or Raven's Standard Progressive Matrices (Raven, Raven, & Court, 2003) that required the same rule combination (see the Appendix). One problem from each pairs of problems was randomly assigned to Set A, and the other remaining problems to Set B. Subjects received each pair of problems in succession, so that they had to complete two problems that required the same rule combination sequentially (see Fig. 3). The order of the problems was counterbalanced so that roughly half of the subjects (n = 103) completed each problem from Set A before they completed the problem's pair from Set B. In the other counterbalanced condition (n = 105), they completed the Set B problems before completing each problem's pair in Set A. The first problem of each pair required a novel rule combination, and the problem immediately following (the other problem in the pair) required the same rule combination. The critical dependent measures from this task were the totals of novel-rule problems and repeated-rule problems answered correctly.

Results

Are novel-rule-combination problems more highly related to WMC and Gf than are repeated-rule-combination problems?

The descriptive statistics for all measures are reported in Table 1, and the complete correlation matrix is presented in Table 2. Composite scores for WMC and Gf were created by calculating the averaged z scores for our WMC measures and

Table 2 Correlation matrix

	OSpan	SymSpan	RunSpan	NovProblems	RepProblems	Letter Sets	NumSeries
OSpan	–						
SymSpan	.51	–					
RunSpan	.48	.46	–				
NovProblems	.18	.32	.38	–			
RepProblems	.33	.41	.48	.53	–		
Letter Sets	.27	.40	.50	.39	.42	–	
NumSeries	.30	.43	.50	.34	.44	.53	–

OSpan = Operation Span; SymSpan = Symmetry Span; RunSpan = Running Span; NovProblems = Novel-Rule-Combination Problems; RepProblems = Repeated-Rule-Combination Problems; NumSeries = Number Series the non-Raven's Gf measures (letter sets and number series). The correlation between the WMC composite and the novel total (r = .36) was statistically lower than the WMC composite's correlation with the repeated total (r = .50), $z(205)$ = −2.29, p = .02 (Steiger, 1980). This finding fails to support the interference/distraction account but does provide support for the opposite conclusion, the learning efficiency account.

We also checked the correlations in both counterbalanced conditions separately, to check whether our results were different across the two counterbalanced conditions. We found that the correlations were higher between the WMC composite and the repeated total (rs = .57 and .45) than between the composite and the novel total (rs = .47 and .27, respectively) in both conditions. Thus, our results were consistent across both counterbalanced conditions. The correlation between the Gf composite and the numbers of novel-rule problems that subjects solved correctly (r = .41) was not significantly different from the Gf composite's correlation with the repeated-rule problems (r = .49), $z(205)$ = −1.30, p = .19.

We also decided to split our sample into high- and low-ability groups by using a median split on the Gf composite, to examine whether restricting the ability range of our subjects would change the relationships between the Raven's problems and WMC. The repeated-rule-combination problems were nominally more highly correlated with WMC than were the novel-rule-combination problems for both the low-ability group (rs = .52 and .39) and the high-ability group (rs = .44 and .24). Thus, our findings are consistent across the entirety of the ability spectrum.

Do novel-rule-combination problems predict unique variance in WMC and Gf above and beyond the repeated-rule-combination problems?

Hierarchical regression analyses were conducted to determine whether performance on the novel-rule-combination problems or repeated-rule-combination problems added incremental validity to the prediction of the WMC composite, above and beyond the other measure. When we first added the repeated total into a model and then the novel total, the change in R^2 was not significant by the traditional cutoff of $p < .05$, ΔR^2 = .013, $F(1, 205)$ =3.72, p = .055. However, when we added the novel total to a model first and then the repeated total, the change in R^2 was significant, ΔR^2 = .131, $F(1, 205)$ = 36.464, $p < .01$.

For predicting the Gf composite, when the repeated total was entered into the first step of a regression model and then the novel total, the change in R^2 was significant, ΔR^2 = .033, $F(1, 205)$ = 9.386, $p < .01$. When the two scores were entered in the reverse order, the change in R^2 was significant as well, ΔR^2 = .101, $F(1, 205)$ = 28.596, $p < .01$.

Discussion

Not only are the results of our study inconsistent with the *interference/distraction* account, they contradict it. One of the reasons that WMC is correlated with Raven's problems is that WMC allows individuals to maintain in memory how they solved previous problems (i.e., the *learning efficiency* account). This is particularly helpful when the next problem that an individual has to solve requires the same rule combination as the previous problem. Even though WMC and Gf are highly related constructs, this study shows a dissociation between the two. Only the WMC relationship to Raven's problems is higher for the repeated-rule-combinations problems. The relationship to other Gf tasks stays the same, regardless of whether the Raven's problem requires a novel or a repeated rule combination.

The particular advantage of our study is that we used a combination of an experimental and a differential approach to research (e.g., Chuderski, 2013; Shipstead & Engle, 2013). Although this type of design requires more subjects in order to counterbalance the order of presentation, it allows for a critical test of the interference/distraction account. We were therefore able to eliminate potential confounds and test whether novel-rule or repeated-rule problems relate more strongly to WMC. We think that this approach will be a very useful one for future individual-differences research. Researchers can use it to essentially treat correlations as a dependent variable and manipulate factors to determine whether those factors change the correlation.

The higher correlations between repeated-rule Raven's problems and WMC in our study are interesting for several reasons. Although attention control plays a large role in the correlation between WMC and Gf (Engle & Kane, 2004), there are many potential reasons for the correlation between WMC and Gf. Researchers have shown that secondary memory is as strongly correlated with Gf as is attentional control (Shipstead et al., 2014; Unsworth & Spillers, 2010). Likewise, one of the reasons that WMC is correlated with Raven's problems is possibly that subjects with high WMC are able to retrieve solutions from previous Raven's problems to solve the current problem. Some researchers have argued that subjects' ability to retrieve information from short-term memory can completely account for WMC's relationship to Gf (Colom et al., 2006; Martínez et al., 2011). Although our results support the learning efficiency hypothesis, we believe that the correlation between Raven's problems and WMC cannot be accounted for by just one mechanism (e.g., dealing with inference or short-term memory). Just as multiple cognitive abilities (i.e., primary memory, secondary memory, and attention control) can account for complex span tasks' intercorrelations, there are multiple reasons why performance on Raven's could be correlated with WMC.

Additionally, the dissociation between WMC and Gf is particularly interesting. Recently, we have theorized that the critical distinction between WMC and Gf is that WMC reflects the ability to maintain easy accessibility of relevant information in memory, and that Gf is related to the ability to disengage from irrelevant information in memory (Shipstead, Harrison, & Engle, 2014). Our findings in the present study support this theory. Only the relationship to WMC is increased for the repeated-rule problems, the ones that would be easier to solve if subjects had the solution of the previous problem still activated in memory.

Author note This work was supported by a grant from the Office of Naval Research (No. N0014-09-1-0129). We thank Kenny Hicks, Thomas Redick, Dakota Lindsey, and Robyn Marshall for their assistance in data collection and for helping with various drafts of this article.

Appendix

Table 3 Specific problems in Raven's mixed

Problem Pair	Set A	Set B	Rule Combination Required
1	30	32	Dist 3, Dist 3, Dist 2
2	22	16	Addition, Subtraction
3	28	34	Dist 3, Dist 3, Dist 3
4	12	E5	Subtraction
5	D8	26	Dist 3, Dist 3

In this table, the problem pairs for the Raven's mixed task are presented. The numbers in Sets A and B denote the problems' numbers in the traditional version of the Raven's Advanced Progressive Matrices. The two problems with letters (i.e., E5 and D8) were taken from Raven's Standard Progressive Matrices and denote the set and number of the problem. The rule combinations required to complete the Raven's advanced problems were taken from Wiley et al. (2011) and, for the Raven's standard problems, were determined by the authors (see Carpenter et al., 1990, for more information on the particular rules). Dist 2 = Distribution of 2; Dist 3 = Distribution of 3

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. doi:10.1037/0033-2909.131.1.30

Broadway, J. M., & Engle, R. W. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods, 42,* 563–570. doi:10.3758/BRM.42.2.563

Bunting, M. (2006). The role of processing difficulty in the predictive utility of working memory span. *Psychonomic Bulletin & Review, 13,* 998–1004. doi:10.3758/BF03193998

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. doi:10.1037/0033-295X.97.3.404

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41,* 244–262.

Colom, R., Abad, F. J., Quiroga, M., Shih, P., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence, 36,* 584–606. doi:10.1016/j.intell.2008.01.002

Colom, R., Rebollo, I., Abad, F. J., & Shih, P. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory & Cognition, 34,* 158–171. doi:10.3758/BF03193395

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

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145–199). New York: Elsevier Science.

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. doi:10.1037/0096-3445.128.3.309

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science, 24,* 2409–2419. doi:10.1177/0956797613492984

Heitz, R. P., Redick, T. S., Hambrick, D. Z., Kane, M. J., Conway, A. R. A., & Engle, R. W. (2006). Working memory, executive function, and general fluid intelligence are not the same. *Behavioral and Brain Sciences, 72,* 135–136.

Jarosz, A. F., & Wiley, J. (2012). Why does working memory capacity predict RAPM performance? A possible role of distraction. *Intelligence, 40,* 427–438. doi:10.1016/j.intell.2012.06.001

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. doi:10.1037/0033-2909.131.1.66

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

Martínez, K., Burgaleta, M., Román, F. J., Escorial, S., Shih, P. C., Quiroga, M., & Colom, R. (2011). Can fluid intelligence be reduced to "simple" short-term storage? *Intelligence, 39,* 473–480.

Matzen, L. E., Benz, Z. O., Dixon, K. R., Posey, J., Kroger, J. K., & Speed, A. E. (2010). Recreating Raven's: Software for systematically generating large numbers of Raven-like matrix problems with normed properties. *Behavior Research Methods, 42,* 525–541. doi:10.3758/BRM.42.2.525

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition, 27,* 759–767. doi:10.3758/BF03198529

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 4. The Advanced Progressive Matrices*. San Antonio: Harcourt Assessment.

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio: Harcourt Assessment.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47,* 609–612. doi:10.1016/j.jrp.2013.05.009

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

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2014). *Working memory capacity and fluid intelligence: Maintenance and disengagement*. Manuscript in preparation

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

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251. doi:10.1037/0033-2909.87.2.245

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

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

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between operation span and Raven. *Intelligence, 33,* 67–81.

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

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

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

Wiley, J., Jarosz, A. F., Cushen, P. J., & Colflesh, G. H. (2011). New rule use drives the relation between working memory capacity and Raven's Advanced Progressive Matrices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 256–263. doi:10.1037/a0021613