Published in final edited form as:

J Cogn Psychol (Hove). 2013 January 1; 25(6): . doi:10.1080/20445911.2013.791299.

Predicting performance on the Raven's Matrices: The roles of associative learning and retrieval efficiency

Lindsey Lilienthal, Elaine Tamez, Joel Myerson, and Sandra Hale
Washington University in St. Louis

Abstract

Previous studies have shown that performance on Williams and Pearlberg's (2006) complex associative learning task is a good predictor of fluid intelligence. This task is similar in structure to that used in studying the fan effect (Anderson, 1974), as both tasks involve forming multiple associations and require retrieval in the face of interference. The purpose of the present study was to investigate the relations among complex associative learning, working memory, and fluid intelligence. Specifically, we asked whether retrieval efficiency, as measured by the fan effect, could account for the relation between complex associative learning and performance on Raven's Advanced Progressive Matrices. Consistent with previous findings, complex associative learning predicted Raven's performance, but the fan effect did not account for this relation. Notably, the learning phase of the fan effect task was significantly correlated with both complex associative learning and Raven's performance, providing further support for the importance of learning as a predictor of fluid intelligence.

Keywords

associative learning; fluid intelligence; fan effect

Fluid intelligence, which refers to the ability to solve novel problems, commonly is measured in adults using Raven's Advanced Progressive Matrices (RAPM; Raven, Raven, & Court, 1998). RAPM is a test of abstract reasoning in which participants are asked to identify the missing piece of a 3×3 array of patterns, presumably by determining which rule or rules govern the patterns found in each row and column. Over the course of the test, participants are exposed to problems on which the rule combinations are increasingly difficult. Many researchers have proposed specific processes or abilities that they believe to be important for successful RAPM performance. For example, Carpenter, Just, and Shell (1990) suggest that the ability to manage problem solving goals and to simultaneously consider multiple rules and answer choices in working memory is important, whereas Wiley, Jarosz, Cushen, and Colflesh (2011) point to the ability to successfully use new combinations of rules and overcome interference from past uses of a particular rule.

Many of the abilities hypothesized to be tapped by RAPM have clear working memory and interference control components, but likely rely on learning as well. For example, before one can simultaneously consider multiple rules in working memory and then switch among them, the rules must be learned. Although much of the previous research with RAPM has focused on the relationship between working memory and fluid intelligence, there has been a revival of interest in learning ability. Recent studies have not only found a strong relation

between fluid intelligence and complex associative learning, but some also suggest that complex associative learning correlates with fluid intelligence as well or better than does working memory (Tamez, Myerson, & Hale, 2008; Williams & Pearlberg, 2006).

In the complex associative learning task pioneered by Williams and Pearlberg (2006), participants learn ten primary words, each of which is associated with three cues (i.e., A, B, and C) that are then each paired with a secondary word (e.g., the primary word LIE is paired with A - fan, B - rim, C - day). At test, participants are given a primary word and a cue (A, B, or C), and must recall the secondary word. Participants are asked to form and maintain associations across four study/test cycles, and the number of correctly recalled secondary words across all four tests is taken as a measure of learning, operationally defined as the ability to accumulate information in memory over multiple exposures. Tamez et al. (2008) reported that complex associative learning was not only related to both working memory and fluid intelligence, but also that associative learning accounted for unique variance in fluid intelligence beyond the contribution of working memory. This result was partially replicated by Kaufman, DeYoung, Gray, Brown, and Mackintosh (2009), who found that both associative learning and working memory predicted unique variance in intelligence.

Although these results are encouraging for researchers interested in understanding the correlates and components of fluid intelligence, the nature of the relation between complex associative learning and fluid intelligence has yet to be fully specified. One possibility is that the associative learning task is measuring individuals' ability to form and maintain associations over time. However, it is also possible that, rather than measuring learning *per se*, the task is measuring individuals' ability to retrieve previously learned information efficiently. The purpose of the present study is to investigate these possibilities and to determine, using correlational and regression analyses, whether complex associative learning is still predictive of performance on RAPM when individual differences in retrieval are also considered.

Retrieval efficiency is often measured by asking participants to verbally recall a supraspan list of words, pictures, or associative pairs (e.g., Shelton, Elliot, Matthews, Hill, & Gouvier, 2010; Unsworth & Spillers, 2010b; Unsworth & Engle, 2007). However, these measures may also reflect differences in encoding that potentially could affect estimates of retrieval. To avoid this problem, in the present study the verification phase of the fan effect task (Anderson, 1974), in which retrieval of over-learned information is measured, was used to measure retrieval efficiency. Importantly, this verification task measures two of the aspects of retrieval (i.e., reactivation of information and interference control) that previously have been found to be important in predicting fluid intelligence (e.g., Unsworth & Spillers, 2010b; Unsworth & Engle, 2007; Bunting, 2006; Lustig, May, & Hasher, 2001). In addition, there are also structural similarities present between the complex associative learning task and the fan effect task.

The fan effect refers to the classic finding that the time required to verify information stored in memory increases with the number of associated items (i.e., fan size) (Anderson, 1974). In the traditional fan effect paradigm, a series of sentences are learned, each one with the structure, "The [person] is in the [place]." Each person typically is associated with one, three, or four places, and each place is also associated with multiple people (i.e., the places are also repetitive, creating overlap between the sentence sets). Individuals are asked to learn several groups of sentences at each fan size. Study/test cycles repeat until the sentences have been successfully learned, at which time a series of both studied and new sentences are presented for verification. It is in this verification stage that the fan effect on retrieval efficiency is observed: On average, it takes more time to verify sentences about people who were associated with a greater number of places. A similar effect has also been found with

error rates in that errors tend to increase with the number of associated places (Anderson & Reder, 1999).

Although other explanations now exist for the fan effect (e.g., the mental model theory, see Radvansky & Zacks, 1991; Radvansky, Spieler, & Zacks, 1993), the first was the Adaptive Control of Thought (ACT) theory (Anderson, 1974), which has since been revised and is now the Adaptive Control of Thought – Rational (ACT-R) theory. Importantly, the ACT-R theory suggests that it is interference, or competition among learned information, that is the cause of the fan effect (Anderson & Reder, 1999; Sohn, Anderson, Reder, & Goode, 2004). This interpretation of the fan effect is directly relevant to the complex associative learning task because this task also requires retrieval in the face of potentially interfering associations. Further, complex working memory span tasks (i.e., working memory tasks that require both the retention and manipulation of information, such as the operation span task) are assumed to require interference control, and differences in the ability to control interference have been postulated to be one important reason for the correlation between working memory and fluid intelligence (e.g. Bunting, Conway, & Heitz, 2004; Engle, Tuholski, Laughlin, & Conway, 1999; Bunting, 2006).

The fan effect has been shown to be quite robust at the group level, and it has been used to explore questions in numerous areas of psychological research, including prospective memory (Cook, Marsh, Hicks, & Martin, 2006) and cognitive aging (Radvansky, Zacks, & Hasher, 1996). However, relatively little research has been done on individual differences in the fan effect. One notable exception is a study by Cantor and Engle (1993), who addressed this issue in the context of a general capacity model of working memory that suggests that limits seen in working memory capacity are primarily the result of limited activation in long-term memory.

To measure individual differences in long-term memory activation, Cantor and Engle (1993) calculated participants' RT slopes from the verification phase of the fan effect task. They reported that the slopes were predictive of verbal ability, as measured by the verbal section of the SAT, and that the slopes accounted for the variance in verbal ability that would have been otherwise attributed to working memory. In addition, Cantor and Engle reported working memory span differences in the fan effect, in that low spans showed an exaggerated effect (i.e., low spans showed a greater increase in RT with increases in fan size than did high spans). This was interpreted as evidence for low span individuals having less activation available to them.

Cantor and Engle (1993) concluded that individual differences in long-term memory activation, as it relates to retrieval ability, were at least partially responsible for the relation between working memory and verbal ability. However, Bunting, Conway, and Heitz (2004) came to a different conclusion regarding working memory and the fan effect, one more consistent with the idea of the importance of controlled attention espoused by Kane, Bleckley, Conway, and Engle (2001). Specifically, Bunting et al. replicated the working memory span differences in the fan effect reported by Cantor and Engle, and went on to demonstrate that these span differences largely disappeared when interference among the learning sets was removed (i.e., each place was only associated with one person, and so was unique). This suggests that what appear to be individual differences in the fan effect due to working memory may also be related to individual differences in the ability to control interference among competing pieces of information. If the associative learning task and the fan effect task are both tapping this aspect of retrieval, this could explain why both predict fluid intelligence.

It is worth noting that the complex associative learning task and the learning phase of the fan effect task have obvious structural similarities. In both tasks, participants are asked to learn items that are each associated with multiple other items. In the associative learning task, these items are unrelated, one-syllable words: A series of ten primary words are each associated with three secondary words. In the fan effect task, these items are the sentences about people in places, with each person associated with between one and four places. The fact that multiple items are associated creates competition and potential interference at retrieval in both tasks. The learning phase of the fan effect task usually receives very little attention, however, because it is viewed as merely a way to ensure that participants have learned the sentences well enough so that when they move on to the verification stage, there will be enough correct responses to assess the fan effect. Instead, it is the RT observations from the verification phase that are thought to measure activation and retrieval ability, and RT has been shown to be related to working memory and verbal ability (Cantor & Engle, 1993).

The primary purpose of the present study is to examine the relations among complex associative learning, the fan effect, and fluid intelligence and investigate to what extent does the fan effect explain, or at least partially explain, the relationship between associative learning and fluid intelligence? If the predictive utility of the complex associative learning task for fluid intelligence is due to retrieval efficiency, then the fan effect task should attenuate the relation between learning and fluid intelligence. However, if complex associative learning is predictive of fluid intelligence because of learning as well as retrieval efficiency, then the fan effect task should not fully attenuate this relation.

Method

Participants

Eighty-six undergraduates (47 female, mean age = 18.96, $SD = 0.98$), all of whom reported English as their primary language, participated in this experiment for course credit. Data from one participant, for whom accuracy on the arithmetic portion of the operation span task was below the eighty percent cut-off, was not included in the analyses. Participants who had previously taken the SAT were asked to self-report their verbal scores, and of the 53 who did so, 96.2 % reported a score at or above 610, compared with 12.5% in the Cantor and Engle (1993) study. Thus, the present sample appears to consist of individuals with much higher verbal ability.

Materials

The experiment consisted of five computerized tasks and a pencil-and-paper Verbal SAT practice test. All computerized tasks were programmed using E-prime 1.2 (Schneider, Eschman, & Zuccolotto, 2002) and were presented to participants using a 17-inch touch screen LCD monitor. Participants made all responses in computerized tasks using a computer mouse, the computer keyboard, or vocally (recorded using an Olympus VN-900PC digital recorder).

Procedure

Participants completed all tasks in one session, lasting approximately 2.5 hours. Participants performed the tasks in the following order: fan effect task, processing speed task, operation span, Verbal SAT practice test, Raven's Advanced Progressive Matrices (RAPM; Raven et al., 1998), and complex associative learning task (Williams & Pearlberg, 2006). Participants were given a break after the operation span task, as well as when needed throughout the experimental session.

Fan effect task—This task was designed based on that used by Cantor and Engle (1993), and the majority of the sentences used in the present study came from those experiments. However, the traditional paradigm only includes fan-size groups of one sentence, three sentences, and four sentences, with two groups at each fan-size level. In the present study, two groups at the fan size of two sentences were added to increase the continuity of the data. Participants were therefore asked to learn a total of twenty sentences, organized into eight groups based on the number of places associated with the person in the sentences. The sentences for the two additional groups were taken from the material used by Bunting et al. (2004). Some of the sentences taken from Cantor and Engle were then adjusted to maintain the original structure of each place being associated with two people.

Participants first were instructed to learn a series of sentences, each with the structure, “The [person] is in the [place]” (e.g., The teacher is in the church.). Participants completed an initial study cycle in which sentences were presented grouped by person. For example, participants saw all four sentences about the teacher on the same screen. The sentence group presentation duration was based on the size of the group, determined using the formula: $n(10s) + 10s$, where n was the number of sentences in the group (Cantor & Engle, 1993).

Once participants had been presented with all eight sentence groups, the initial recall stage began. Participants were shown the name of a person (e.g., teacher) at the top of the screen and were asked to type the sentences they remembered about that person one at a time into a text box. After typing one sentence, they pressed the Enter key and were then shown two boxes, one labeled “More” and one labeled “Done.” If they remembered additional sentences about that person, participants were instructed to click “More,” and if not, to click “Done” to move on to the next person. Participants recalled all sentences for each group in this way for the duration of the task. Participants were not told at recall how many sentences there were for that person, and they did not receive feedback.

Following the initial recall test, participants were presented with the same sentences one at a time and in a random order, for twenty seconds per sentence. Participants then were asked to again recall the sentences in the same way as was previously described. This cycle of learn-test was repeated until participants recalled all of the sentences in each sentence group correctly three times (performance on the initial recall test was not included in this score). As soon as this criterion was met for a sentence group, it dropped out of the cycle. Once participants reached the criterion for all sentence groups, they moved on to the final study phase, which was identical to the initial study phase (i.e., sentences were again presented in their respective groups, organized by person, rather than individually).

In the final phase of the task, the verification phase, participants were shown a series of both old (studied) and new (foil) sentences one at a time and were asked to respond as quickly and accurately as possible to the question, “Is this a sentence you studied?” Participants indicated their decision by pressing one of two keys on the keyboard. Foil sentences were largely the same as those used in Cantor and Engle (1993), and included the same people and places as the studied sentences but were paired differently to create novel sentences. Where changes were made to accommodate the new sentence groups, the same pattern of novel pairing was used. The measure that is usually of primary interest in this task is the RT to the studied and the foil sentences during this final verification phase; the number of cycles required to reach the criterion for all sentences during the acquisition phase was also recorded as a measure of learning ability.

Categorization RT task—Participants were presented with words one at a time on the computer screen. Each word was either the name of an animal (e.g., bear) or a fruit/vegetable (e.g., lime). Participants were asked to respond to the question of whether or not

each word was the name of an animal by pressing one of two keys on the keyboard. As soon as a response was made, participants were shown the next word. Participants were instructed to respond as quickly and accurately as possible, and RT and accuracy were recorded.

Operation span—In this verbal complex working memory task (Turner & Engle, 1989), participants were instructed to remember a series of words. Before the presentation of each word, participants were shown an arithmetic equation (e.g., $(4 \times 2) - 1 = 5$). They were instructed to read each equation out loud and decide whether the given solution was correct, indicating this decision by pressing one of two keys on the keyboard. At the end of each series, participants were prompted to recall all the words from that series out loud and in the exact order of presentation. The number of words per series ranged from two to seven, and participants completed two test trials of each length for a total of twelve trials. The series lengths were presented in a random order that was the same for all participants. Prior to the test trials, participants completed six practice trials consisting solely of equations and four practice trials that included both equations and words. A proportion correct scoring procedure was used, in which participants were awarded one point for each word recalled in the correct serial position and one half of one point for each word recalled in the incorrect serial position. Proportions were averaged across the twelve trials for a measure of working memory capacity.

Verbal SAT practice test—In order to assess verbal ability, participants were asked to complete one verbal section of a SAT practice test (Section 2 of Practice Test 2; Robinson, Katzman, et al., 2009). This section was made up of eight multiple-choice sentence-completion questions (e.g., After the governor's third trip overseas, voters complained that he was paying too little attention to ---- affairs. A. intellectual, B. extraneous, C. specialized, D. aesthetic, E. domestic), four reading comprehension questions based on two short passages, and twelve reading comprehension questions based on one long passage. Participants had 25 minutes to answer as many of the 24 questions as possible. All participants completed all 24 questions in the allotted time. The questions used in the practice tests were taken from SAT tests from 1981-1984 and 1992. Performance was measured as the proportion of questions answered correctly.

Raven's Advanced Progressive Matrices (RAPM)—This task was a computerized version of the RAPM (Raven et al., 1998). Following Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle (2004), only the odd items from the second set of RAPM problems were administered, and participants were given ten minutes to complete as many of the eighteen items as possible to save time and prevent fatigue or frustration. On each trial, participants saw a 3×3 matrix of patterns with the lower right hand pattern removed. Participants were given eight possible solutions and were asked to choose the pattern that best completed the matrix. Participants were also given a "Do Not Know" option. Performance was measured as the number of problems answered correctly out of eighteen.

Complex associative learning task—In this task (Williams & Pearlberg, 2006; Tamez et al., 2008), participants were instructed to learn associations between a primary word (e.g., lie) and three secondary words (e.g., fan, rim, day). Participants first saw the primary word, and then were prompted to press the "A" key. After pressing the "A" key, participants saw an associated secondary word, and were instructed to press "Enter" when they were ready. Participants were then instructed to press the "B" key and the "C" key, with each action resulting in the presentation of an associated secondary word. Participants saw ten primary words, each of which was associated with three secondary words, one for each of the three cues (i.e., A, B, and C), for a total of thirty word associations. After all thirty associations were presented, participants were tested. In the test trials, a primary word and one of the

three cues (i.e., A, B, or C) appeared, and participants were asked to type the corresponding secondary word into a text box. Participants were tested over all associations, and then moved on to a second learning phase. The primary words were presented in a random order at test, and participants were prompted to recall all three secondary words for that primary word, in the A-B-C order, before moving on to another primary word. This learn-test cycle was repeated a total of four times. All trials were self-paced, and accuracy during the tests was recorded. The proportion of correct trials was averaged across the four tests and was used as a measure of learning ability.

Results

Descriptive statistics for performance on all six tasks are presented in Table 1. Both accuracy and RT in the verification stage of the fan effect task were analyzed using 2 (trial type: target, foil) $\times$ 4 (fan size: 1, 2, 3, 4) repeated measures ANOVA. With respect to accuracy, there was no main effect of trial type and no interaction between fan size and trial type, both F s < 1.0 . Although accuracy was greater than 95%, on average, for all fan sizes, there was a main effect of fan size, $F(3, 252) = 4.66$, $p = .003$, partial $\eta^2 = .053$, reflecting the fact that accuracy decreased as a function of fan size, consistent with the fan effect (see Figure 1; Anderson & Reder, 1999).

With respect to RT, the repeated measures ANOVA revealed main effects of both fan size, $F(3, 252) = 16.76$, $p < .001$, partial $\eta^2 = .166$, and trial type, $F(1, 84) = 86.76$, $p < .001$, partial $\eta^2 = .508$, but no size $\times$ type interaction, $F(3, 252) = 1.88$, *ns*. Although RT on foil sentences were longer on average than RT on target sentences, RT increased with fan size to a similar degree for both target and foil sentences (see Figure 2). Thus, the fan effect was also observed in RT at the group level, for both target and foil sentences.

At the individual level, RTs were strongly correlated with fan size (mean $r = .59$, median $r = .64$), evidence that the fan effect also exists at an individual level. Therefore, following Cantor and Engle (1993), a response time slope was calculated for each participant. Unlike in Cantor and Engle (1993), slopes were not correlated with operation span or Verbal SAT, nor were they correlated with performance on RAPM. Significant correlations were found between both the RT and accuracy from the fan effect task and other measures, suggesting that although in the present study the slopes appear to be unreliable, other measures taken from the fan effect task do not share this problem.

With respect to the complex associative learning task, a one-way repeated measures ANOVA revealed a significant effect of test block, $F(3, 252) = 385.17$, $p < .001$, partial $\eta^2 = .821$, reflecting the fact that performance improved systematically across the four blocks (mean percent correct on the first test: 25.3, $SD = 21.4$; mean percent correct on the fourth test: 79.6, $SD = 21.3$). Following Williams and Pearlberg (2006; see also Tamez et al., 2008), the average number of items correct across all four tests was taken as a measure of learning (i.e., how much information could be accumulated in secondary memory given repeated exposure and retrieval practice), for the purpose of individual differences analyses.

Before conducting individual difference analyses, further steps were taken to normalize all RT distributions at the group level. Because group RT distributions are typically positively skewed, a log transformation was applied to the RT data from the fan effect task as well as from the processing speed task. Additionally, because no interaction between fan size and trial type was observed for either RT or accuracy, for both measures data from target and foil trials were combined. Finally, in order to most effectively capture the fan effect while minimizing the number of correlations, only data from the smallest and largest fan sizes were included in the analyses.

The correlations among the various measures are presented in Table 2. Performance on RAPM was correlated significantly with both measures of learning (associative learning and fan cycles) as well as with operation span. Performance on RAPM was also correlated significantly with fan-size-four RT and accuracy, but not with processing speed, fan-size-one RT, or fan-size-one accuracy. Verbal SAT performance was significantly correlated with fan cycles, operation span, and RAPM. Finally, the two learning measures were significantly correlated, as were all RT measures. Interestingly, no significant correlations were found between the RT and accuracy measures at either fan size, suggesting the two measures are independent assessments of performance on the fan effect task.

In order to examine the relations among performance on RAPM, learning, and retrieval efficiency as measured by the fan effect, composite variables were constructed when possible, using a mean Z-score approach. Composites are advantageous because they shift the focus toward constructs and away from specific measures, tasks or tests, increase reliability, and also reduce type one error by reducing the number of comparisons. Three composites were computed: A Learning composite was created from the associative learning task and the fan cycles learning measure, a Fan-Size-One composite was created from the RT and accuracy measures on fan-size-one trials, and a Fan-Size-Four composite was created from the RT and accuracy measures on fan-size-four trials. The fan cycles, fan-size-one RT, and fan-size-four RT measures were multiplied by negative one prior to being standardized and entered into the composite, so that better performance always was indicated by a higher number.

The fan composites were formed using both RT and accuracy information because both showed the fan effect, and although on fan-size-four trials both RT and accuracy were significantly correlated with performance on RAPM, RT and accuracy were themselves uncorrelated, suggesting that they represent independent assessments of the fan effect (see Table 2). In the following regression analyses, the Fan-Size-One composite will always be entered into the model before the Fan-Size-Four composite, statistically controlling for baseline speed and accuracy and thus making the Fan-Size-Four composite a measure of the fan effect.

Hierarchical regression analyses with performance on RAPM as the dependent variable are presented in Table 3. The magnitude of the fan effect (i.e., the Fan-Size-Four composite after controlling for the Fan-Size-One composite) was a significant predictor of RAPM performance, accounting for 12% of the RAPM variance (Model 1). When the Learning composite and the fan effect were both entered as predictors, the total RAPM variance accounted for increased to 18% (Models 2 and 3). Importantly, the Learning composite predicted unique RAPM variance beyond that accounted for by the fan effect (Model 2). Similarly, the fan effect predicted unique RAPM variance when it was entered into the model after the Learning composite (Model 3). Finally, when working memory was entered into the model after the Learning composite and the fan effect, it uniquely accounted for an additional 5% of RAPM variance (Model 4). In contrast, when working memory was entered first followed by the fan effect, adding the Learning composite did not significantly increase the RAPM variance accounted for (Model 5).

Hierarchical regression analyses with performance on the Verbal SAT practice test as the dependent variable are presented in Table 4. The fan effect was a significant predictor of Verbal SAT performance, although the total amount of variance accounted for was relatively low at only 5% (Model 6). Working memory added significantly to the model when entered after the fan effect, and when the fan effect was entered after working memory, it accounted for only a marginally significant amount of variance in Verbal SAT practice test performance (Models 7 and 8). The Learning composite, when entered after working

memory and the fan effect, did not contribute significantly to the model (Model 9). When the Learning composite was entered first followed by the fan effect and then working memory, the contributions of Learning and working memory were both only marginally significantly (both $ps < .06$; Model 10).

Discussion

The primary purpose of the present study was to determine whether the ability to retrieve information from secondary memory in the face of competing associations can explain why associative learning predicts fluid intelligence. Accordingly, the present study examined the relations among performance on RAPM, the fan effect, and performance on two learning tasks (the complex associative learning task and the learning phase of the fan effect task). Complex associative learning was correlated with RAPM, as was a new measure of associative learning, fan cycles (i.e., cycles to criterion in the learning phase of the fan effect task). RT and accuracy on fan-size-four trials were also correlated with RAPM performance, as well as with complex associative learning. However, when the fan effect (i.e., the Fan-Size-Four composite, controlling for the Fan-Size-One composite) and the Learning composite were entered in hierarchical regression, each made independent contributions to predicting performance on RAPM. Thus, both retrieval efficiency in the face of competing associations and the ability to form and accumulate associations over time are important independent predictors of fluid intelligence.

The present finding that complex associative learning predicts performance on RAPM is consistent with other recent studies investigating associative learning (e.g., Kaufman et al., 2009; Tamez et al., 2008; Williams & Pearlberg, 2006). Tamez et al. (2008) found that complex associative learning was able to account for unique RAPM variance beyond the contribution of two measures of working memory (operation span and grid span). Kaufman et al. (2009) used structural equation modeling to investigate the relations among these variables, measuring associative learning ability using both the complex associative learning task and a paired associates learning task. Kaufman et al. found that associative learning accounted for unique variance in RAPM, consistent with Tamez et al., but also that working memory accounted for unique variance in RAPM as well. In the present study, working memory accounted for the same RAPM-related variance as complex associative learning, and learning did not account for unique variance in RAPM.

Although the present study and those of Tamez and Kaufman differ with respect to the relations among learning, working memory, and fluid intelligence, all three studies indicate that associative learning is an important predictor of performance on measures of intelligence. Additionally, complex associative learning appears to be an important factor in the relation between age and fluid intelligence in adults: Tamez, Myerson, and Hale (2012) found that a cascade model that included paths from age to speed, speed to working memory (and also directly to fluid intelligence), from working memory to learning, and from learning to fluid intelligence accounted for 79% of the variance in fluid intelligence. In contrast, a similar cascade model that did not include learning accounted for only 49% of the variance, suggesting that associative learning is also an important factor in cognitive aging.

The major finding of the present study was that retrieval efficiency in the face of competing associations, as measured by the fan effect, did not explain the relation between complex associative learning and fluid intelligence. Although the present study found significant fan effects for both RT and accuracy, Figure 1 clearly shows that a linear increase in RT is only observed on fan sizes one through three. The leveling off observed between fan sizes three and four is not predicted by any fan effect theory, but it is not unusual in the literature. For example, similar fan effects were reported by Bunting et al. (2004) in their Experiments 1A

and 1B: Although RT increased significantly from fan-size-one trials to fan-size-three trials, there was no difference between sizes three and four. Importantly, unlike RT, overall accuracy did not level off in the present study, as the most errors were made with fan-size-four sentences (Table 1). Accordingly, Fan-Size-One and Fan-Size-Four composites created in the present study took both RT and accuracy into account.

Inhibition and the ability to control interference are central to some theories of working memory (e.g., Kane et al., 2001), and have been implicated as important factors in the relation between working memory and fluid intelligence (e.g., Wiley et al., 2011; Bunting, 2006). It therefore seemed possible that interference control was also the important link between complex associative learning and fluid intelligence, but the present results did not support this idea. Both the fan effect and associative learning were significant predictors of RAPM performance and shared much of the accounted for variance, but associative learning also accounted for a significant amount of unique RAPM variance. This indicates that the two associative learning measures used in the present study tap some aspect of retrieval efficiency similar to what is measured by the fan effect; however, this is not the whole story. That is, the present results suggest that fluid intelligence is related not only to retrieval efficiency and interference control, but also to learning ability.

The present study's finding that retrieval ability cannot fully account for the relation between complex associative learning and performance on the RAPM suggests that the unique predictive power of associative learning may have reflected individual differences in encoding. Whether such encoding differences were because of differences in encoding strategy is not known, as participants in the present study were not asked to report their strategy use; however, previous studies have reported that individual differences in strategy use can be important in episodic memory performance. For example, Bailey, Dunlosky, and Kane (2008) found that effective strategy use on working memory tasks partially mediated the relation between working memory span and episodic recall. In addition, Unsworth and Spillers (2010a) found that individuals with high working memory spans outperformed individuals with low working memory spans on episodic memory tasks that emphasized both strategic encoding and contextual retrieval, suggesting that both encoding and retrieval play an important role in the relation between working memory capacity and episodic recall. Unsworth and Spillers' results are consistent with what was observed in the present study, although the complexity of the complex associative learning task likely made strategy differences especially important in the present study.

Future work investigating the exact nature of the unique predictive power of complex associative learning is needed. Such efforts are likely to be more fruitful if they include multiple measures of each construct, thereby permitting the use of a structural equation modeling (SEM) approach that examines the relation between error-free estimates of the latent constructs involved. One possible source of error in the present study was the use of the same order of task administration for all participants. This approach has long been used in studies of individual differences (e.g., Cantor & Engle, 1993; Horn & Cattell, 1966; Salthouse & Pink, 2008; Vernon & Jensen, 1984) because counterbalancing can introduce noise that reduces the observed correlations among measures. The use of a single order requires the assumption that order effects such as fatigue and practice will not affect the ranking of individuals on a task. This approach is not without flaws, however, in large part because the underlying assumption is not usually tested. Using SEM obviates these concerns, allowing one to use counterbalancing without affecting the correlations between latent constructs, and therefore would be the preferred strategy for future work, given an adequate sample size.

The present study is the first to investigate whether the fan effect is predictive of fluid intelligence, but two previous studies have explored the relations among the fan effect, working memory, and verbal ability. Both Cantor and Engle (1993) and Bunting et al. (2004) reported that the fan effect was related to working memory capacity, in that low span individuals showed an exaggerated fan effect. Cantor and Engle (1993) also reported that the fan effect was able to explain the relation between working memory and Verbal SAT performance. Because both of these previous studies included individuals who varied greatly in cognitive ability, of interest in the present study was whether similar results would be obtained in a high ability sample.

The present results show that although both working memory and the fan effect both were related to performance on a Verbal SAT practice test in our high ability sample, working memory was still able to account for unique variance beyond that contributed by the fan effect (i.e., the relation between working memory and Verbal SAT was not fully explained by the fan effect). In addition, performance on the working memory measure, operation span, was not significantly correlated with any measure of the fan effect (i.e., fan-size-four RT, fan-size-four accuracy, or the fan effect). This suggests, then, that there are differences between ability groups. It is possible that these differences could reflect the relative restriction of range in the present study, but this possibility seems less likely when one considers the correlations observed among other measures, namely among performance on RAPM, working memory, and both measures of associative learning. Indeed, the correlation of .35 observed between operation span and RAPM is similar to that typically observed with more diverse samples (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). Alternatively, others have proposed that relations among cognitive abilities may differ across ability groups. For example, Spearman's law of diminishing returns posits that correlations among cognitive abilities are lower in individuals with high IQ (Jensen, 2003; Spearman, 1927). This could explain the discrepancies between the results of the present study and those reported by Cantor and Engle. However, evidence regarding the validity of Spearman's law of diminishing returns has been mixed (e.g., Reynolds, Keith, & Beretvas, 2010; Hartmann & Teasdale, 2004; Deary et al., 1996), and these issues deserve more attention.

Previous studies have typically ignored the number of fan cycles as a measure from the learning phase of the fan effect task, and focused instead on RT and accuracy data from the verification phase. Interestingly, however, fan cycles in the present study was a valuable measure of learning, correlating significantly with complex associative learning, working memory, Verbal SAT, and RAPM. The two measures of learning actually provide complementary measures in that the complex associative learning task is a measure of how many associations one can learn in a specified number of study/test cycles, whereas fan cycles is a measure of how many such cycles it takes for one to learn a specified number of associations. Importantly, even though there are large differences between complex associative learning and fan cycles in both the materials used in the tasks (unrelated one syllable words vs. sentences with overlapping people and places) and the measures themselves (number of associations learned vs. number of cycles required to learn a set of associations), the two measures were correlated with RAPM to a similar magnitude. Thus, even when associative learning is measured in very different ways, there appears to be a robust relation between the construct of associative learning and performance on RAPM.

The relation among fan cycles and the RT and accuracy measures from fan-size-one and fan-size-four trials in the present study is also worth noting. In the learning phase of the fan effect task, participants completed as many study/test cycles as necessary for them to learn the material to criterion (i.e., until they recalled each sentence correctly three times). As in previous studies, this criterion was meant to ensure that the material had been sufficiently

learned and to minimize individual differences in encoding, thereby making the following verification stage purer measure of retrieval efficiency. In the present study, fan cycles was not significantly correlated with either fan-size-one RT or fan-size-four RT. However, this was not the case for verification accuracy at either fan size, as both were significantly correlated with fan cycles. Although everyone was taken to the same criterion in the learning phase, faster learners (i.e., those who reached criterion in less cycles) were more accurate during the verification stage than were slower learners on both fan-size-one and fan-size-four sentences. Importantly, this also suggests that verification accuracy is not only influenced by the number of exposures one has to the material, as poor learners by definition are those who both studied and were tested over the material more times. Thus, learning ability on the fan effect task is an important contributor to verification accuracy, and its role is not eliminated by having everyone reach the same recall criterion.

The complex associative learning task and the fan effect task are structurally similar in that both ask participants to learn and maintain multiple associations over time and to resolve potential interference among items in the learning sets. Despite their similarities, the results of the present study do not support the idea that the complex associative learning task is predictive of RAPM performance simply because the learning task measures retrieval efficiency in the face of competing associations. Complex associative learning and the fan effect both accounted for significant unique variance in RAPM performance, suggesting that each measure is tapping something different and that both learning and retrieval efficiency are important aspects of fluid intelligence.

Acknowledgments

The efforts of Lindsey Lilienthal and Elaine Tamez were supported by NIA Training Grant AG00030.

References

- Anderson JR. Retrieval of propositional information from long-term memory. *Cognitive Psychology*. 1974; 6:451–474.
- Anderson JR, Reder LM. The fan effect: New results and new theories. *Journal of Experimental Psychology: General*. 1999; 128:186–197.
- Bailey H, Dunlosky J, Kane MJ. Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory and Cognition*. 2008; 36:1383–1390. [PubMed: 19015498]
- Bunting M. Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2006; 32:183–196.
- Bunting MF, Conway ARA, Heitz RP. Individual differences in the fan effect and working memory capacity. *Journal of Memory and Language*. 2004; 51:604–622.
- Cantor J, Engle RW. Working-memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1993; 19:1101–1114.
- Carpenter PA, Just MA, Shell P. What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review*. 1990; 97:404–431. [PubMed: 2381998]
- Conway ARA, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*. 2005; 12:769–786. [PubMed: 16523997]
- Cook GI, Marsh RL, Hicks JL, Martin BA. Fan effects in event-based prospective memory. *Memory*. 2006; 14:890–900. [PubMed: 16938699]
- Deary IJ, Egan V, Gibson GJ, Austin EJ, Brand CR, Kellaghan T. Intelligence and the differentiation hypothesis. *Intelligence*. 1996; 23:105–132.

- Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*. 1999; 128:309–331. [PubMed: 10513398]
- Hartmann P, Teasdale TW. A test of Spearman's 'law of diminishing returns' in two large samples of Danish military draftees. *Intelligence*. 2004; 32:499–508.
- Horn JL, Cattell RB. Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*. 1966; 57:253–270. [PubMed: 5918295]
- Jensen AR. Regularities in Spearman's law of diminishing returns. *Intelligence*. 2003; 31:95–105.
- Kane MJ, Bleckley MK, Conway ARA, Engle RW. A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*. 2001; 130:169–183. [PubMed: 11409097]
- Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RW. The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General*. 2004; 133:189–217. [PubMed: 15149250]
- Kaufman SB, DeYoung CG, Gray JR, Brown J, Macintosh N. Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*. 2009; 37:374–382.
- Lustig C, May CP, Hasher L. Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*. 2001; 130:199–207. [PubMed: 11409099]
- Radvansky GA, Spieler DH, Zacks RT. Mental model organization. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 1993; 19:95–114.
- Radvansky GA, Zacks RT. Mental models and fact retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 1991; 17:940–953.
- Radvansky GA, Zacks RT, Hasher L. Fact retrieval in younger and older adults: The role of mental models. *Psychology and Aging*. 1996; 11:258–271. [PubMed: 8795054]
- Raven, J.; Raven, JC.; Court, JH. *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. Oxford, England: Oxford Psychologists Press; 1998.
- Reynolds MR, Keith TZ, Beretvas SN. Use of factor mixture modeling to capture Spearman's law of diminishing returns. *Intelligence*. 2010; 38:231–241.
- Robinson, A.; Katzman, J. Staff of the Princeton Review. *Cracking the SAT, 2010 Edition*. New York, NY: Random House; 2009. Practice Test 2, Section 2.
- Salthouse TA, Pink JE. Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review*. 2008; 15:364–371. [PubMed: 18488653]
- Schneider, W.; Eschman, A.; Zuccolotto, A. *E-Prime users guide*. Pittsburgh, PA: Psychology Software Tools; 2002.
- Shelton JT, Elliott EM, Matthews RA, Hill BD, Gouvier WD. The relationships of working memory, secondary memory, and general fluid intelligence: Working memory is special. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2010; 36:813–820.
- Sohn MH, Anderson JR, Reder LM, Goode A. Differential fan effect and attentional focus. *Psychonomic Bulletin and Review*. 2004; 11:729–734. [PubMed: 15581125]
- Spearman, CE. *The abilities of man*. London: Macmillan; 1927.
- Tamez E, Myerson J, Hale S. Learning, working memory, and intelligence revisited. *Behavioural Processes*. 2008; 78:240–245. [PubMed: 18308483]
- Tamez E, Myerson J, Hale S. Contributions of associative learning to age and individual differences in fluid intelligence. *Intelligence*. 2012.1016/j.intell.2012.04.004
- Turner ML, Engle RW. Is working memory capacity task dependent? *Journal of Memory and Language*. 1989; 28:127–154.
- Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*. 2007; 114:104–132. [PubMed: 17227183]

- Unsworth N, Spillers GJ. Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*. 2010a; 17:200–205. [PubMed: 20382920]
- Unsworth N, Spillers GJ. Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*. 2010b; 62:392–406.
- Vernon PA, Jensen AR. Individual and group differences in intelligence and speed of information processing. *Personality and Individual Differences*. 1984; 5:411–423.
- Wiley J, Jarosz AF, Cushen PJ, Colflesh GJH. New rule use drives the relation between working memory capacity and Raven's Advanced Progressive Matrices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2011; 37:256–263.
- Williams BA, Pearlberg SL. Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence*. 2006; 34:177–191.

Figure 1. Mean error rates as a function of fan size for both target and foil trial types. Data are from the verification phase of the fan effect task.

Figure 2. Mean response time (in milliseconds) as a function of fan size for both target and foil trial types. Data are from the verification phase of the fan effect task.

Table 1

Descriptive Statistics for All Tasks

Task	M	SD	Range	Skew	Kurtosis
Fan Cycles	4.54	1.20	3 – 8	0.66	-0.02
Fan Size 1 RT	1443	414	877 – 2885	1.47	2.27
Fan Size 2 RT	1610	449	1036 – 3946	2.37	8.74
Fan Size 3 RT	1720	469	1075 – 3483	1.29	1.87
Fan Size 4 RT	1688	437	1088 – 3315	1.47	2.78
Fan Size 1 Accuracy	.98	.04	.83 – 1.0	-2.00	3.51
Fan Size 2 Accuracy	.96	.06	.71 – 1.0	-2.52	7.53
Fan Size 3 Accuracy	.97	.04	.81 – 1.0	-1.68	3.34
Fan Size 4 Accuracy	.95	.04	.77 – 1.0	-1.58	3.87
Associative Learning	.57	.21	.17 – .96	-0.06	-0.85
RAPM	.65	.15	.27 – 1.0	-0.07	-0.08
Verbal SAT	.79	.13	.33 – 1.0	-1.16	1.47
Operation Span	.58	.14	.26 – .90	-0.04	-0.42
Categorization RT	609	68	495 – 787	0.27	-0.88

Note. Fan Size RT and Categorization RT data are in milliseconds. Fan Size Accuracy, Associative Learning, RAPM, Verbal SAT, and Operation Span are measured as the proportion of items correct.

Table 2

Correlations

Variable	1	2	3	4	5	6	7	8	9
1. Categorization RT	-								
2. Fan Size 1 RT	.36	-							
3. Fan Size 4 RT	.37	.79	-						
4. Fan Size 1 Accuracy	.25	.07	.09	-					
5. Fan Size 4 Accuracy	.12	.01	.01	.13	-				
6. Fan Cycles	-.04	.07	-.05	-.29	-.35	-			
7. Associative Learning	.03	-.21	-.28	.12	.20	-.33	-		
8. Operation Span	-.10	-.15	-.17	.09	.10	-.23	.54	-	
9. Verbal SAT	-.22	-.13	-.11	-.11	.18	-.25	.09	.28	-
10. RAPM	-.09	-.13	-.22	.00	.27	-.29	.23	.35	.31

Note. Significant correlations are in bold. Correlations .22 and .27 are significant at $p < .05$; correlations .28 are significant at $p < .01$.

Table 3

Summary of Hierarchical Regression Analysis for Variables Predicting Fluid Intelligence

Predictor	R ²	R ²	F(R ²)	df
<i>Model 1</i>				
1. Fan Size 1 Composite	.01	.01	0.74	1, 83
2. Fan Size 4 Composite	.13	.12	10.95 ^{**}	1, 82
<i>Model 2</i>				
1. Fan Size 1 Composite	.01	.01	0.74	1, 83
2. Fan Size 4 Composite	.13	.12	10.95 ^{**}	1, 82
3. Learning Composite	.18	.05	5.37 [*]	1, 81
<i>Model 3</i>				
1. Learning Composite	.10	.10	9.65 ^{**}	1, 83
2. Fan Size 1 Composite	.10	.00	0.01	1, 82
3. Fan Size 4 Composite	.18	.08	7.48 ^{**}	1, 81
<i>Model 4</i>				
1. Learning Composite	.10	.10	9.65 ^{**}	1, 83
2. Fan Size 1 Composite	.10	.00	0.01	1, 82
3. Fan Size 4 Composite	.18	.08	7.48 ^{**}	1, 81
4. Working Memory	.23	.05	5.17 [*]	1, 80
<i>Model 5</i>				
1. Working Memory	.13	.13	12.00 ^{**}	1, 83
2. Fan Size 1 Composite	.13	.00	0.09	1, 82
3. Fan Size 4 Composite	.22	.09	9.22 ^{**}	1, 81
4. Learning Composite	.23	.01	1.37	1, 80

Note.^{*}
 $p < .05$,^{**}
 $p < .01$.

Table 4

Summary of Hierarchical Regression Analysis for Variables Predicting Verbal SAT

Predictor	R ²	R ²	F(R ²)	df
<i>Model 6</i>				
1. Fan Size 1 Composite	.00	.00	0.03	1, 83
2. Fan Size 4 Composite	.05	.05	4.14 *	1, 82
<i>Model 7</i>				
1. Fan Size 1 Composite	.00	.00	0.03	1, 83
2. Fan Size 4 Composite	.05	.05	4.14 *	1, 82
3. Working Memory	.11	.06	5.82 *	1, 81
<i>Model 8</i>				
1. Working Memory	.08	.08	6.97 *	1, 83
2. Fan Size 1 Composite	.08	.00	0.01	1, 82
3. Fan Size 4 Composite	.11	.03	3.08	1, 81
<i>Model 9</i>				
1. Working Memory	.08	.08	6.97 *	1, 83
2. Fan Size 1 Composite	.08	.00	0.01	1, 82
3. Fan Size 4 Composite	.11	.03	3.08	1, 81
4. Learning Composite	.12	.00	0.35	1, 80
<i>Model 10</i>				
1. Learning Composite	.04	.04	3.77	1, 83
2. Fan Size 1 Composite	.05	.00	0.20	1, 82
3. Fan Size 4 Composite	.08	.03	2.62	1, 81
4. Working Memory	.12	.04	3.67	1, 80

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