Why does working memory capacity predict RAPM performance? A possible role of distraction

Andrew F. Jarosz *, Jennifer Wiley
Department of Psychology, University of Illinois at Chicago, United States

Article info
Article history:
Received 2 July 2011
Received in revised form 7 June 2012
Accepted 9 June 2012
Available online 26 June 2012

Keywords:
Intelligence
Working memory capacity
Attentional control

ABSTRACT
Current theories concerning individual differences in working memory capacity (WMC) suggest that WMC reflects the ability to control the focus of attention and resist interference and distraction. The current set of experiments tested whether susceptibility to distraction is partially responsible for the established relationship between performance on complex span tasks and the Raven's Advanced Progressive Matrices (RAPM). This hypothesis was examined by manipulating the level of distraction among the incorrect responses contained in RAPM problems, by varying whether the response bank included the most commonly selected incorrect response. When entered hierarchically into a regression predicting a composite score on span tasks, items with highly distracting incorrect answers significantly improved the predictive power of a model predicting an individual's WMC, compared to the model containing only items with less distracting incorrect responses. Additional analyses were performed examining the types of errors that were made. A second experiment used eye-tracking to demonstrate that these effects seem to be rooted in differences in susceptibility to distraction as well as strategy differences between high and low WMC individuals. Results are discussed in terms of current theories about the role of attentional control in performance on general fluid intelligence tasks.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction
Since the earliest days of psychology, one recurring topic has been the measurement of intelligence. The idea that there may be stable differences between individuals in their general abilities that allow some to learn faster and more easily than others, particularly in educational settings, is indeed an intriguing concept. As such, a considerable amount of individual differences research has focused on the study of intelligence, with one interesting finding being the correlation in performance on tests of working memory capacity (WMC) and general fluid intelligence (gF; Kane et al., 2004). In particular, the Raven's Advanced Progressive Matrices (RAPM, Raven, Raven, & Court, 1998) has been used to investigate the WMC–gF relationship in numerous experiments and has consistently been shown to correlate with performance on complex span tasks at around .30 (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Kane et al., 2004; Unsworth & Engle, 2005; Wiley & Jarosz, 2012; Wiley, Jarosz, Cushen, & Colflesh, 2011). While this correlation has been reliably demonstrated, there are still many questions about it that are unanswered. In particular, what is it about tests of WMC that predicts performance on the RAPM?

In practice, WMC is usually measured by complex span tasks that assess the ability to hold multiple objects in memory while performing a concurrent processing task. As shown in the left side of Fig. 1, the operation span (Ospan) task involves remembering lists of words while simultaneously verifying math equations. Ospan presents processing and memory components in each item, followed by a test for the memory components at the end of the set of items. In contrast, the
RAPM (Raven et al., 1998) was originally designed as a test of the ability to find meaning in complex stimuli. The RAPM has been found to load heavily on to measures of gF across numerous studies (e.g., Kane et al., 2004; Marshalek, Lohman, & Snow, 1983), and is considered a prototypical test of gF by the designers of the test (Raven et al., 1998). Each item consists of a 3×3 matrix of figures that change along both rows and columns according to certain rules, with the bottom right figure missing. The problem solver is instructed to look both along the rows and down the columns in order to select the figure that correctly completes the pattern from a bank of potential responses beneath the matrix. The right side of Fig. 1 depicts a RAPM-like problem where the correct answer is response option 5, which can be reached by following a progression rule along the rows (adding a dot to each consecutive figure) as well as a distribution of three rule (each shape appearing once in each row/column). In the RAPM, test items are arranged in terms of their normed difficulty, with the easiest problems presented first and the hardest problems presented last.

While WMC tasks and the RAPM share relatively few surface features, the relationship between the two has been repeatedly demonstrated, and has resulted in numerous suggestions about why they are related. One early idea suggested that it is the number of rules and goals that must be stored in memory that drives the relationship between the two (Carpenter, Just, & Shell, 1990; Mulholland, Pellegrino, & Glaser, 1980). That is, the rules governing the progression of figures in the RAPM item's matrix, and the goals of the problem solver, must be held in WMC while completing each item. This account suggests that as items become more difficult over the course of the RAPM and require more rules and goals to solve, the relationship between the RAPM and WMC should increase. This capacity account is in line with some of the current capacity theories of WMC, such as the time-based/resource sharing model of WMC (Barrouillet, Bernardin, & Camos, 2004). Essentially, a high WMC individual has more resources to store information while concurrently processing, and as such can better keep track of goals and rules in a given item as items become more difficult.

An additional approach is a learning account (Guthke & Stein, 1996; Verguts & De Boeck, 2002a, 2002b). Having demonstrated that repeated use of a set of rules in an RAPM-like task increases the likelihood of applying those rules to later items (Verguts & De Boeck, 2002b) and improves solution rates compared to sets of items requiring different rules to solve (Verguts & De Boeck, 2002a), Verguts and de Boeck suggested that WMC may relate to the ability to learn complex rule combinations. Unfortunately, Verguts and de Boeck either did not measure WMC in their experiments (Verguts & De Boeck, 2002b), or when they did, they failed to analyze the relation between WMC and performance on the interleaved vs. repeated blocks of rules (Verguts & De Boeck, 2002a). However, according to a learning account, high WMC should lead to an increased ability to learn rules, and
one would expect to see high WMC individuals benefit from repeated experience with the same rules, while low WMC individuals would show less benefit.

Another possibility is that WMC and the RAPM are related because both depend on the ability to control one’s attention. According to this attentional control account, as the RAPM progresses there is proactive interference from previously encountered items, and high WMC individuals are better able to resist this interference and generate solutions to the new items. This is in line with both the idea that complex tasks such as intelligence tests may require increased executive control (Marshalek, Lohman, & Snow, 1983) and that performance on WMC tasks is related to the ability to control attention and resist interference from distracting or previously encountered stimuli (Engle, 2002; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2003; Unsworth & Engle, 2007).

An increasing amount of evidence has accumulated in support of the attentional control account. As stated earlier, if a capacity account was correct, the relationship between WMC and RAPM solution should increase as the test progresses, due to an increasing reliance on WMC as the number of rules, rule tokens, and goals in the problems increases. However, Salthouse (1993) found a relatively consistent correlation between WMC and performance on individual RAPM items across item order. Likewise, if the capacity account was correct, one would predict an increase in the correlations between WMC and RAPM solution as the number of rule tokens (i.e., instances of a rule) increased. Unsworth and Engle (2005) tested this, demonstrating that not only were the point-biserial correlations between performance on each item in the RAPM and WMC relatively constant across item order, but also that the relationship did not depend on the number of rule tokens in the item. If anything, items with one rule token showed the strongest relationship to WMC. Indeed, Verguts and De Boeck (2002a) demonstrated that even if one attempts to minimize the number of rules required to solve matrix completion items, their solution rate still correlates with WMC. Clearly, the capacity account fails to explain the WMC–RAPM relationship.

Similarly, recent data is difficult to explain via a learning account. Wiley, Jarosz, Cushen, and Colflesh (2011) demonstrated that items containing the first presentation of a combination of rules drove the WMC–RAPM correlation, with items that repeated a previous rule having a significantly lower correlation with WMC. These results are in direct contradiction to the learning account, which would have predicted the repeated rule items to have a stronger correlation with WMC. Further, additional analyses failed to support the capacity account, as the correlation between WMC and RAPM performance did not increase with problem order or number of rules. Wiley et al. suggested that these results are most consistent with an attentional control account, with the ability to direct one’s attention to new combinations of rules driving the correlation between WMC and the RAPM.

Thus, the attentional control account seems to provide a better explanation of the WMC–RAPM relationship than either the capacity or learning accounts. Further, this account seems consistent with recent suggestions that WMC may fundamentally be related to attentional control, or the ability to control one’s focus of attention (Engle, 2002; Kane et al., 2001; Kane & Engle, 2003). If high WMC individuals are better able to control their focus of attention, it would allow them to resist the influence of previously learned rule combinations, and enable them to combine rules in novel ways or find new rules in order to solve problems. Conversely, low WMC individuals may be unable to resist interference from previous solution attempts, and their performance suffers (Wiley et al., 2011). Thus, it may well be that it is the general ability to resist interference and distraction that underlies individual differences in performance on both complex span tasks and the RAPM.

Continuing with this theme, the present studies investigate whether distraction from incorrect, yet salient, potential responses may be another factor influencing the RAPM’s correlation with WMC. In the standard RAPM, each item contains eight potential solutions. It is entirely possible that the presence of eight possible responses diverts attention away from solution processes, exposing test takers to salient distractors. This is supported by eye-tracking research suggesting that poorer solvers spend more time examining the response bank (Vigneau, Caissie, & Bors, 2006), although that study did not also examine WMC. As such, it is possible that low WMC individuals’ performance may suffer because they are unable to ignore or resist interference from salient distractors within the response bank.

2. Experiment 1

The present study tests this distraction hypothesis by manipulating the presence of the most frequently chosen incorrect responses within the response bank of RAPM items. So that no new response options needed to be created, the number of responses in the response bank was reduced from eight possible responses to four (as shown on the left side of Fig. 2). In the high salience versions of the problems, the most commonly selected incorrect response was included among the four response choices (as determined by norming data from the RAPM manual; Raven et al., 1998). For example, suppose the most common incorrect response in the example item (Fig. 1) was response option 2. In the high salience version of the item, response option 2 from the original item appeared as one of the incorrect responses, as seen on the left side of Fig. 2. On the right side of Fig. 2, the most common incorrect response is replaced by another incorrect response, making this the low salience version of the item. The presence of the salient distractor or its control was randomized across possible response positions.

Each participant received items in both conditions, assigned to even and odd problem sets, counterbalanced across participants. Based on the assumption that the most common incorrect response is also the most salient incorrect response, manipulating its presence should increase or decrease the distractibility of the response bank. Thus, if distraction due to salient incorrect responses accounts for part of the variance between WMC and the RAPM, then the correlation between WMC and RAPM performance should be stronger on the high salience problems compared to the low salience problems.

2.1. Error analysis

In addition to manipulating the presence of highly salient incorrect responses, one can also consider qualitative differences among error types. The RAPM manual has assigned the most common error on each item into one of four categories.
Incomplete solutions, arbitrary lines of reasoning, over-determined choices, and repetitions. Incomplete solution errors involve selecting an answer that is only partially correct, due to a failure to account for all of the elements of the figures within an item. For example, in Fig. 1, response choice 2 would be an example of an incomplete solution error, as it involves leaving out a “progression” rule from the correct solution. Arbitrary-line-of-reasoning errors stem from the solver using a reasoning process other than the one required by the item (e.g., using incorrect or inappropriate rules to attempt solving the problem). This could be seen in instances where the solver inappropriately applies a rule from a previous item on a problem at hand, or generates an improper rule for the item. Response choice 3 could reflect an arbitrary-line-of-reasoning error in Fig. 1, as one would have to apply a subtraction rule across the bottom row in the problem matrix, rather than a progression rule, in order to reach that answer. Over-determined choice errors involve selecting an answer that contains as many elements of the figures as possible (seen in Fig. 1, response choice 7, which contains all of the elements of the figures contained in the item), while repetition errors involve simply choosing a response that mimics one of the original figures in the matrix (seen in Fig. 1, response option 4, which is a repetition of the item just above the blank cell in the matrix). It is possible to make predictions regarding how the two most common error types will correlate with WMC. Because so few items result in over-determined choice and repetition errors, and because of the fact that these error types would be unlikely to differentiate between theories, the primary focus of these predictions will be on incomplete solution and arbitrary-line-of-reasoning errors. To begin with, a capacity account would suggest that selection of incomplete solution responses would have a strong correlation with WMC, due to low WMC individuals’ inability to keep track of all the rules and elements required to solve a given item. Encountering an incorrect option containing most of the elements and transformations in the item would therefore be more likely to mislead the low WMC individual. On the other hand, one would not expect selection of arbitrary-line-of-reasoning responses to correlate with WMC according to a capacity account, as the capacity account makes no suggestions regarding errors in reasoning, unless those are based upon the failure to include information from the problem due to a lack of capacity (as would be seen in items with incomplete solution errors).

A learning account would predict that selection of responses stemming from arbitrary-line-of-reasoning errors should relate to WMC. An arbitrary-line-of-reasoning response may activate a previously learned rule that is incorrect for the current problem, particularly if that rule was used recently. Given previous evidence that repeated use of rules can lead to a mental set for those rules (Verguts & De Boeck, 2002b), an interesting prediction is that the learning account may actually suggest an increase in the selection of these types of distractors as WMC increases. That is, as individuals learn rules they may be more likely to be misled by a distracter related to those rules in a later problem. Therefore, a learning account would predict an increase in the selection of arbitrary-line-of-reasoning responses as WMC increases, resulting in poorer problem solving performance and, thus, a negative correlation between WMC and performance on problems where the distracter stems from an arbitrary line of reasoning. Incomplete solution responses would be unlikely to tap into any previously learned rules, making their selection unlikely to relate to WMC according to the learning account.

Finally, an attentional control account, like the learning account, makes no predictions regarding capacity, and as such would not predict that the selection of incomplete solution responses would relate to WMC. On the other hand, an attentional control account predicts that low WMC individuals would be more likely to be affected by proactive interference (Kane & Engle, 2000). If responses representing arbitrary lines of reasoning activate previously learned rules, low WMC individuals may be more likely to perseverate in applying those rules. Therefore, an attentional control account would predict an increase in the selection of responses stemming from arbitrary lines of reasoning for low WMC individuals, resulting in decreased problem accuracy. Because high WMC individuals are more resistant to interference, there should be a positive correlation between WMC and performance on problems where the distracter stems from an arbitrary line of reasoning.

In summary, two predictions can be made based on the attentional control account. First, if distraction due to salient incorrect responses partially accounts for the correlation between WMC and the RAPM, then performance on high-salience items will have a stronger correlation with WMC than performance on low salience items, and will be a better overall predictor of WMC. Second, if WMC reduces interference from previously encountered rule combinations, then performance on items where the salient distracter results from

Fig. 2. Example high salience (left) and low salience (right) versions of items. Note that the only difference is in the presence of the salient distracter (option 2 in the high salience version). For both versions of the item, response option 1 is the correct answer.
an arbitrary line of reasoning should be a better predictor of WMC than performance on items with distracters representing incomplete solutions.

2.2. Method

2.2.1. Participants

Sixty-four undergraduate students (23 male) enrolled in Introduction to Psychology were recruited from the subject pool at the University of Illinois at Chicago.

2.3. Materials

2.3.1. RAPM

The RAPM was presented via computer, with the standard 40 minute time limit (Raven et al., 1998). All 36 items in Set II were preceded by 2 practice problems drawn from the beginning of Set I. Two versions of each item from the original RAPM task were created. The high salience version contained the most commonly selected incorrect response in the response bank, while the low salience version did not. In each case the response bank was reduced to four items so that new materials did not need to be created (specific details of which original responses were used, and the order in which they were presented, are included in the Appendix). High salience and low salience conditions were assigned to either even or odd items and were counterbalanced across subjects. This allowed for a within-subjects manipulation. RAPM items are presented in order of normative difficulty, making the manipulation of alternating items a good way to balance difficulty between problem conditions. An analysis of previously collected data demonstrated that even and odd RAPM items do not differ in their correlation with WMC, $t(252) = .60$, ns.

Participants were instructed to select the figure from the response bank that correctly completed the matrix both along the row and down the column.

2.3.2. Complex span tasks

Two complex span tasks were used as measures of WMC: the reading span (Rspan) and the Ospan. Fig. 1 contains examples of a typical trial for each of these span tasks. Both were presented via computer with answers written down on paper, following the standard administration procedure as outlined in Conway et al. (2005). The complex span tasks present alternating processing and memory components within each trial, followed by a test for the memory components at the end of the trial. For each item in the Rspan, participants are requested to read a sentence out loud, judge whether or not the sentence makes sense, and then remember a letter. Between two and five items are presented one at a time with experimenter-controlled timing to eliminate rehearsal between items. After the last item in the trial, participants are cued to write down all the letters they can remember in the correct serial position. Three trials of each length are presented over the course of the test. The Ospan task follows the same procedure, however the processing component requires deciding whether a math equation is correct or incorrect, and the memory component requires remembering a word. The two tasks are not timed, but generally take 15 min each to complete.

2.3.3. Paper folding task

Both Part 1 and Part 2 of the paper folding task (Kit of Factor-Referenced Cognitive Tasks; Ekstrom, French, Harman, & Dermen, 1976), were presented using paper and pencil with a 3 minute time limit on each part of the test (see Fig. 3 for an example item). This task involves two sets of 10 items (Part 1 and Part 2), with instructions to determine what a folded, hole-punched piece of paper would look like if it were unfolded. Paper folding is considered a test of gF shown to correlate with the standard RAPM (Kane et al., 2004). Since the RAPM is being altered for the purposes of this study, paper folding provides a measure of gF that can be used to test the correlation between high and low salience items with an unaltered test of gF.

2.3.4. Card rotation task

The card rotation task (Kit of Factor-Referenced Cognitive Tasks; Ekstrom et al., 1976) was administered using paper and pencil with a 3 minute time limit on each of Parts 1 and 2 of the test (see Fig. 3 for an example item). This task contains two sets of 10 target figures (Part 1 and Part 2). Each target figure is followed by eight additional figures. Instructions require determining whether each additional figure is simply rotated and otherwise the same as the target, or whether it has undergone some additional manipulation (such as flipping it). Card rotation is a spatial ability test that is not related to gF, to be used to partial out variance due to spatial ability from the paper folding task, allowing for a better assessment of gF.

2.3.5. Procedure

Consent was obtained and participants first completed the RAPM, with test version counterbalanced between subjects. Following this, participants completed the two WMC tasks, with task order again counterbalanced between subjects. Participants then completed the paper folding task, followed by the card rotation task. The entire session took approximately 90 min to complete.

2.3.6. Scoring

The RAPM was scored by taking the proportion of items correctly solved. Separate scores were created for high and low salience items. Additionally, subscores on the RAPM problems were developed for each participant based on the type of error the salient distracter in each item stemmed from, as reported in the RAPM manual (Raven et al., 1998). Two subscores were created, using the average score of participants on items with incomplete solution and arbitrary line of reasoning distracters.

The complex span tasks were scored by taking the average proportion of memory items correctly recalled in each trial. All individuals completed at least 80% of the processing components of trials correctly, confirming that individuals engaged in both parts of the complex span tasks. Scores on the two span tasks correlated at .59, which is typical of correlations found in research using span tasks (Conway et al., 2005). Span task scores were combined into a composite score by transforming them into $Z$ scores, then averaging the resulting variables.

The paper folding task was scored by taking the proportion of items answered correctly. The card rotation task was scored by taking the proportion of figures correctly identified as being the same or different from the target.
2.4. Results

2.4.1. Descriptive statistics and reliability

Descriptive statistics for all tasks are shown in Table 1. Correlations for all measures are presented in Table 2. Due to time constraints, only 45 participants completed the card rotation task, and only 60 participants completed the paper folding task. All participants completed all other tasks. Reliability for tasks was computed using parallel form reliability for the complex span tasks, after splitting span data into three forms for each test. Each form contained one trial of each trial length. All other measures use Cronbach’s alpha for reliability.

Because major alterations were performed on the RAPM items to create the two versions of the test, performance on the high and low salience versions of each item was compared to a previously collected data set of 255 individuals from the same university who completed the traditional RAPM during previous semesters (see Fig. 4). Overall, performance was better on the current version of the RAPM with only 4 response options when compared to the original, t(317) = −6.04, p < .001. However, the shapes of the curves remained similar.

Even with the changes in the RAPM measure, WMC and RAPM scores were significantly correlated, r = .40, p < .001. Likewise, the partial correlation of the new RAPM to paper folding, controlling for spatial ability by covarying card rotation performance, was significant, r = .43, p = .004.

Table 1
Descriptive statistics for measures.

<table>
<thead>
<tr>
<th>Task</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original RAPM</td>
<td>255</td>
<td>0.55</td>
<td>0.16</td>
<td>0.17</td>
<td>0.97</td>
<td>.80</td>
</tr>
<tr>
<td>Low salience RAPM</td>
<td>64</td>
<td>0.71</td>
<td>0.17</td>
<td>0.33</td>
<td>1</td>
<td>.83</td>
</tr>
<tr>
<td>High salience RAPM</td>
<td>64</td>
<td>0.65</td>
<td>0.17</td>
<td>0.22</td>
<td>1</td>
<td>.68</td>
</tr>
<tr>
<td>Ospan</td>
<td>64</td>
<td>0.63</td>
<td>0.15</td>
<td>0.34</td>
<td>1</td>
<td>.85</td>
</tr>
<tr>
<td>Rspan</td>
<td>64</td>
<td>0.68</td>
<td>0.15</td>
<td>0.37</td>
<td>0.93</td>
<td>.75</td>
</tr>
<tr>
<td>Paper folding</td>
<td>60</td>
<td>0.66</td>
<td>0.21</td>
<td>0.15</td>
<td>1</td>
<td>.83</td>
</tr>
<tr>
<td>Card rotation</td>
<td>45</td>
<td>0.91</td>
<td>0.11</td>
<td>0.49</td>
<td>1</td>
<td>.89</td>
</tr>
<tr>
<td>Incomplete solution</td>
<td>64</td>
<td>0.62</td>
<td>0.21</td>
<td>0.00</td>
<td>1.00</td>
<td>–</td>
</tr>
<tr>
<td>Arbitrary reasoning</td>
<td>64</td>
<td>0.57</td>
<td>0.22</td>
<td>0.00</td>
<td>1.00</td>
<td>–</td>
</tr>
</tbody>
</table>

Note: Original RAPM data is from 255 students from the same population, collected from a previous study. Because odd and even items were completed by different participants, RAPM Cronbach’s alphas (α) for salience manipulated items are represented with the alpha for odd items first, followed by the alpha for even items.

2.4.2. Effects of the salience manipulation

There was an overall effect for problem type, with participants performing better overall on low salience problems, t(63) = 3.30, p = .002, d = .83 (see Table 1 for descriptive statistics of measures). For each item, the proportion of the total incorrect responses represented by the manipulated response choice (i.e., the salient distracter in the high salience condition) was calculated. The manipulated response choice was significantly more likely to be selected in the high salience condition (M = .51, SD = .32) than in the low salience condition (M = .26, SD = .22), t(35) = 4.21, p < .001.

As shown in the left panel of Fig. 5, high salience problems (r = .46, p < .001) had a stronger correlation with the composite span score than low salience problems (r = .21, p = .096; t(61) = 2.28, p = .026). To further explore this relationship, a hierarchical regression predicting composite span score was performed with low salience item accuracy as a predictor in the first step, and high salience item accuracy as a predictor in the second step. While the initial model was marginally significant, F(1, 62) = 2.86, p = .10, the addition of high salience item performance resulted in a significant model, with a significant change in the R² value, R² = .46, ΔR² = .17, ΔF(1, 61) = 12.79, p = .001. In the final model, low salience item performance did not predict composite span score (β = −.06, t(61) = −.46, ns), while high salience item performance did (β = .49, t(61) = 3.58, p = .001, sr² = .41). Repeating the analysis with high salience item performance entered first followed by low salience item performance in the second step resulted in no significant change in the R² value, R² = .46, ΔR² = .003, ΔF(1, 61) = .21, ns.

A second set of analyses predicted paper folding performance after controlling for spatial ability to see if high salience items were also a better predictor of gF than low salience items. A hierarchical regression predicting paper folding performance with card rotation performance entered in the first step as a control for spatial ability demonstrated a significant improvement in model fit with the addition of high salience and low salience problem accuracy in the second step, R = .72, R² change = .12, F(2, 41) = 5.09, MSE = .022, p = .01. Both card rotation performance (β = .479, t(41) = 4.00, p < .001, sr² = .435) and high salience problem performance (β = .292, t(41) = 2.36, p = .02, sr² = .257) predicted paper folding performance, while low salience problem performance (β = .142, t(41) = 1.12, ns) did not predict any unique variance.
In sum, performance on high salience items had a stronger correlation with WMC when compared with low salience item performance, high salience item performance explained variance in WMC beyond that explained by low salience item performance, and only performance on high salience items uniquely predicted WMC. Similarly, in a regression predicting performance, and only performance on high salience item performance explained variance in WMC beyond that explained by low salience item performance, high salience item performance explained variance in WMC beyond that explained by low salience item performance, and remained the only unique predictor. This follows the prediction of the attentional control account, suggesting that high WMC individuals are better able to avoid distraction from the highly salient incorrect option within the response bank. This result is in line with previous research suggesting that WMC aids in resisting interference from visually distracting stimuli (Darowski, Helder, Zacks, Hasher, & Hambrick, 2008; Sanchez & Wiley, 2006).

Converging evidence was obtained from an analysis of items by the type of salient distractor they contained. While performance on both incomplete solution and arbitrary-line-of-reasoning items was correlated with the composite score of-reasoning items entered in the second step. The initial model was significant, \( F(1, 62) = 4.62, p = .035 \), with incomplete solution item performance predicting composite span scores, (\( β = .26, t(62) = 2.15, p = .04, r^2 = .26 \)). However, the addition of arbitrary line of reasoning item performance in the second step resulted in a significant change in the \( R^2 \) value, \( R = .46, ΔR^2 = .14, ΔF(1, 61) = 11.16, p = .001 \). In the final model, incomplete solution item performance did not predict composite span scores (\( β = .03, t(61) = .25, ns \)), while arbitrary line of reasoning item performance did (\( β = .44, t(61) = 3.34, p = .001, r^2 = .38 \)). Repeating the analysis with arbitrary line of reasoning item performance entered first followed by incomplete solution item performance in the second step resulted in no significant change in the \( R^2 \) value, \( R = .46, ΔR^2 = .001, ΔF(1, 61) = .06, ns \).

### 2.5. Discussion

The results of this study strongly support the idea that salient distracters among response options contribute to the WMC–RAPM correlation. While performance on both high and low salience items remained correlated with the composite score of the complex span tasks, performance on high salience items had a stronger correlation. More importantly, when placed hierarchically into a regression, performance on the high salience items predicted variance in the composite span score above and beyond low salience item performance, and remained the only unique predictor. This follows the prediction of the attentional control account, suggesting that high WMC individuals are better able to avoid distraction from the highly salient incorrect option within the response bank. This result is in line with previous research suggesting that WMC aids in resisting interference from visually distracting stimuli (Darowski, Helder, Zacks, Hasher, & Hambrick, 2008; Sanchez & Wiley, 2006).

Converging evidence was obtained from an analysis of items by the type of salient distracter they contained. While performance on both incomplete solution and arbitrary-line-of-reasoning items was correlated with the composite score of-reasoning items entered in the second step. The initial model was significant, \( F(1, 62) = 4.62, p = .035 \), with incomplete solution item performance predicting composite span scores, (\( β = .26, t(62) = 2.15, p = .04, r^2 = .26 \)). However, the addition of arbitrary line of reasoning item performance in the second step resulted in a significant change in the \( R^2 \) value, \( R = .46, ΔR^2 = .14, ΔF(1, 61) = 11.16, p = .001 \). In the final model, incomplete solution item performance did not predict composite span scores (\( β = .03, t(61) = .25, ns \)), while arbitrary line of reasoning item performance did (\( β = .44, t(61) = 3.34, p = .001, r^2 = .38 \)). Repeating the analysis with arbitrary line of reasoning item performance entered first followed by incomplete solution item performance in the second step resulted in no significant change in the \( R^2 \) value, \( R = .46, ΔR^2 = .001, ΔF(1, 61) = .06, ns \).
of the complex span tasks, performance on arbitrary-line-of-reasoning items had a marginally stronger correlation. When placed into a regression, performance on arbitrary-line-of-reasoning items predicted composite span scores beyond performance on incomplete solution items, and remained the only unique predictor. The fact that only arbitrary-line-of-reasoning distracters predicted WMC is also consistent with the attentional control account.

An interesting question about these findings is the extent to which they are the result of different strategy use among low and high WMC individuals. Early work by Bethel-Fox, Lohman, and Snow (1984) suggested two primary strategies for solving visual analogy problems: constructive matching and response elimination. A constructive matching strategy involves looking at the problem, figuring out what the answer should look like, and then searching for that solution. For the RAPM, this would entail looking at the problem matrix, formulating what the correct response should look like, and then going on to seek out the correct response in the response bank. In contrast, a response elimination strategy entails going through each response in the response bank, while low WMC individuals may resort to less effortful strategies (e.g., a constructive matching strategy) that involve looking at the problem, figuring out what the correct response should look like, seeing if it fits as a correct answer to the problem, and if not, eliminating that response as an option and moving on to the next one. Later work (Vigneau et al., 2006) confirmed that better solvers of the RAPM seemed to favor a constructive matching strategy, while worse solvers seemed to favor a response elimination strategy. Given previous work indicating that high and low WMC individuals can differ in their approaches to problem solving (Beilock & DeCaro, 2007; Wiley & Jarosz, 2012), one might argue that the distraction effects observed in Experiment 1 may stem from different solution strategies being employed by high and low WMC individuals. That is, high WMC individuals may have the resources necessary to implement more effortful strategies (e.g., a constructive matching strategy) that involve creating and maintaining a representation of what the correct answer should look like before searching for the solution in the response bank, while low WMC individuals may resort to less taxing strategies (such as response elimination) that involve checking each potential response against the other figures in the matrix. As a by-product of this strategy, they end up getting more exposure to the potential distracter.

The use of different strategies could also be seen as consistent with the results of the error analyses. Rejecting responses representing arbitrary lines of reasoning may pose a greater challenge for those using a response elimination strategy. With a constructive matching strategy, one already has a response in mind before exposure to the response bank, potentially inoculating the individual against a distracter based on an incorrect set of solution rules. However, a person using a response elimination strategy is not necessarily approaching items with a given set of rules in mind, and as such may be more likely to have incorrect rules activated by the distracting response. Thus, to the extent that low WMC individuals are more inclined to use a response elimination strategy, they may be more likely to expose themselves to incorrect responses and have incorrect rules primed by distracters. The investigation of this hypothesis is the main goal for Experiment 2.

3. Experiment 2

One way to assess the role of strategy use as a possible cause for the effects of salient distracters is through an eye-tracking paradigm. By analyzing the eye movements of individuals solving the modified RAPM, it should be possible to see whether high and low WMC individuals differ in their inspection patterns, including time spent on the salient distracters. As mentioned above, previous research has identified two common strategies for solving visual analogy problems: a constructive matching strategy and a response elimination strategy (Bethel-Fox et al., 1984). These strategies have also been applied to the RAPM (Vigneau et al., 2006), with eye-tracking evidence suggesting that more successful solvers on the RAPM tend towards the constructive matching strategy, while less successful solvers are more likely to utilize the response elimination strategy. Specifically, Vigneau et al. found that those more successful on a shortened version of the RAPM spent a larger proportion of their time on the problem matrix, and a smaller proportion of time on the response options. Similarly, more successful participants toggled between the problem matrix and response area less often than less successful participants.

If strategy differences are found between high and low WMC individuals on all problems, similar to those found by Vigneau et al. (2006) for more and less successful solvers, this would suggest that susceptibility to distraction is being driven by a response-elimination strategy. In contrast, if low WMC solvers are observed to spend a disproportionate amount of time looking at the salient distracter or response bank particularly in the high salience condition, this would suggest that low WMC individuals’ performance is being affected by their susceptibility to distraction over and above any strategic differences that might exist. To test between these possible
three measures were used to look for strategy differences between high and low WMC individuals: the rate of toggling between the response options and the problem matrix, the proportion of overall time spent on the response bank, and the proportion of overall time spent looking at the salient distracters. The number of toggles was calculated by summing the number of times that participants looked from the problem matrix to the response bank and from the response bank back to the problem matrix. Rates were determined using proportions of total time. Descriptive statistics for all eye movement data are presented in Table 4.

3.3.3. Rate of toggling between problem and response bank

A hierarchical regression with toggle rate on low salience items predicting composite span score in the first step and toggle rate on high salience items in the second step was performed to look for strategy differences due to WMC. Toggle rate for low salience items was significant in the initial model, \( R = .46, F(1, 33) = 9.03, p = .01 \). However, adding toggle rate for high salience items in the second step resulted in a significant change in the \( R^2 \) value, \( R = .58, \Delta R^2 = .12, \Delta F(1, 32) = 5.57, p = .03 \). In the final model, toggle rate for low salience items did not predict composite span score (\( \beta = .30, t(32) = .85, ns \)), while toggle rate on high salience items negatively predicted composite WMC (\( \beta = -.84, t(32) = -2.36, p = .03, sr^2 = -.34 \)). This suggests that those with higher WMC toggled less often between the problem matrix and response bank, compared to those with low WMC, and that this effect was greater on high salience items. Repeating the analysis with high salience item performance entered first followed by low salience item performance in the second step resulted in no significant change in the \( R^2 \) value, \( R = .53, \Delta R^2 = .006, \Delta F(1, 32) = .25, ns \). These results replicate the findings of Experiment 1.

3.3.2. Eye movement analyses

Based on the work of Vigneau et al. (2006), three measures were used to look for strategy differences between high and low WMC individuals: the rate of toggling between the response options and the problem matrix, the proportion of overall time spent on the response bank, and the proportion of overall time spent looking at the salient distracters. The number of toggles was calculated by summing the number of times that participants looked from the problem matrix to the response bank and from the response bank back to the problem matrix. Rates were determined using proportions of total time. Descriptive statistics for all eye movement data are presented in Table 4.

<table>
<thead>
<tr>
<th>Task</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low salience RAPM</td>
<td>35</td>
<td>0.67</td>
<td>0.14</td>
<td>.39</td>
<td>.94</td>
<td>.66, .51</td>
</tr>
<tr>
<td>High salience RAPM</td>
<td>35</td>
<td>0.64</td>
<td>0.18</td>
<td>.22</td>
<td>.94</td>
<td>.65, .74</td>
</tr>
<tr>
<td>Ospan</td>
<td>35</td>
<td>0.66</td>
<td>0.17</td>
<td>.41</td>
<td>1.00</td>
<td>.78</td>
</tr>
<tr>
<td>Sspan</td>
<td>35</td>
<td>0.52</td>
<td>0.20</td>
<td>.22</td>
<td>.93</td>
<td>.79</td>
</tr>
</tbody>
</table>

Note: Because odd and even items were completed by different participants, RAPM Cronbach’s alphas (\( \alpha \)) for salience manipulated items are represented with the alpha for odd items first, followed by the alpha for even items.
performance in the second step resulted in no significant change in the R² value, R = .58, ΔR² = .02, ΔF(1, 32) = .72, ns.

3.3.4. Proportion of time on response bank
To further examine strategy differences due to WMC, a hierarchical regression predicting composite WMC scores was performed with the proportion of time spent looking at the response bank in low salience items entered in the first step, and high salience items entered in the second step. While the initial model was significant, R² = .38, F(1, 33) = 5.57, p = .02, the addition of the proportion of time on the response bank for high salience items in the second step resulted in a significant change in the R² value, R² = .52, ΔR² = .12, ΔF(1, 32) = 5.41, p = .03. In the final model, proportion of time on the response bank for low salience items did not predict composite span score, (β = .05, t(32) = .19, ns) while proportion of time on the response bank for high salience items was a significant, negative predictor (β = −.55, t(32) = −2.33, p = .03, sr² = −.35) suggesting that high WMC individuals spent less time on the response bank particularly on high salience items compared to low WMC individuals. Repeating the analysis with proportion of time on the response bank for high salience items entered first resulted in no significant change in the R² value, R² = .52, ΔR² = .001, ΔF(1, 32) = .04, ns.

Because the rate of toggling between the problem matrix and the response bank on the low salience items can be seen as a measure of the default solution strategy that individuals may employ, an additional hierarchical regression was performed including this measure in the first step, to see if strategies might fully account for differences in time spent on the response bank. In this additional analysis, the rate of toggling on low salience items was entered into the first step of the regression, with the proportion of time spent looking at the response bank on low salience items entered into the regression in the second step, and high salience items entered into the third step. As before, toggle rate for low salience items was significant in the initial model, R = .46, F(1, 33) = 9.03, p = .01. Adding the proportion of time on the response bank for low salience items did not result in a significant change in the model, ΔF(1, 32) = .03, ns. However, the addition of the time on the response bank for high salience items resulted in a marginal increase in the R² value, R² = .55, ΔR² = .09, ΔF(1, 31) = 3.86, p = .06. In the final model, neither the rate of toggling on low salience items (β = −.31, t(31) = −1.23, ns) nor the time on the response bank for low salience items (β = .23, t(31) = .82, ns) predicted composite span scores, while proportion of time on the response bank for high salience items was a marginal, negative predictor (β = −.48, t(31) = −1.96, p = .06, sr² = −.30). This suggests that, even controlling for differences in the rate of toggling between individuals, high WMC individuals were still less likely to look at the response bank when a salient distracter was present. Repeating the analysis with measures entered in the reverse order resulted in no significant change in the R² value, even with the addition of the rate of toggling on low salience items, R = .55, ΔR² = .03, ΔF(1, 31) = 1.52, ns.

3.3.5. Proportion of time on salient distracters
Finally, a regression with the proportion of time spent looking at the salient distracters predicting composite WMC was used to look for differences in susceptibility to distraction. This analysis could only be performed on the data from the high salience items. The overall model was significant, R² = .46, F(1, 33) = 8.79, p = .01. Proportion of time spent looking at the salient distracters was negatively related to WMC, β = −.46, t(33) = −2.96, p = .01, sr² = −.46.

3.4. Discussion
The above results replicate Experiment 1 in demonstrating that performance on high salience items predicts WMC over and above performance on low salience items. Because effects were seen most strongly in the high-salience conditions, these data suggest that low WM participants are more susceptible to distraction from highly salient incorrect responses. In both the regression looking at the rate of toggling between the problem matrix and response bank and the regression examining the proportion of time spent looking at the response bank, high salience items predicted variance above and beyond low salience items, while low salience items did not do the same. This suggests that the salient distracters in particular are attracting the attention of the low WMC solvers. Likewise, high WMC individuals spent less time on the salient distracters than low WMC individuals, again suggesting that the high WMC participants are less susceptible to distraction by the highly salient incorrect responses.

That said, some effects were seen even on the low salience items. The greater rate of toggling between the response bank and the problem matrix among low WMC participants on the low salience problems, as well as their increased time looking at the response bank on low salience items, suggests that low WMC individuals are more likely to use a response elimination strategy than high WMC individuals. Thus, the present results suggest, rather than a pure susceptibility-to-distraction explanation, that the relationship between WMC and performance on RAPM items with a salient distracter may be exacerbated by different tendencies in strategy use. However, it is worth noting that, even when accounting for strategy-related variance by including the rate of toggling and time on response bank for low salience items, time on the response bank for high salience items continued to predict additional variance in WMC. Strategies therefore seem to only partially account for the effect of salient distracters on the WMC–RAPM relationship.

4. General discussion
The overall results of the two experiments make several suggestions about the relation of WMC to RAPM performance. First, it is clear that the presence of a salient distracter differentially affects individuals based on their WMC, with
those lower in WMC being more negatively impacted by salient distracters than those with high WMC. However, rather than being purely due to differences in susceptibility to distraction, these effects appear to be partially influenced by strategy use. Higher WMC individuals seem to favor a constructive matching strategy to a greater extent than lower WMC individuals. This result is evident in the fact that low WMC individuals toggle between the problem matrix and the response bank more often than high WMC individuals even on low-saliency distracter problems. This is in line with previous research suggesting that high WMC individuals may default to more complex strategies than low WMC individuals (Beilock & DeCaro, 2007; Wiley & Jarosz, 2012). However, while strategy differences may have played a role, the additional variance explained by high salience items in these analyses, as well as the increased time spent on the salient distracter by low WMC individuals, suggest that the effects of the salient distracter are also due to it capturing the attention of the problem solver. Indeed, even accounting for strategy-related variance did not completely eliminate the impact of the distracter on the relationship between WMC and the RAPM. Thus, these data suggest that WMC contributes to better RAPM performance in at least in two ways: first, by allowing high WMC individuals to use strategies that limit exposure to potentially distracting responses, and second, by providing resistance to distraction from incorrect responses that they are exposed to.

These findings make sense in the greater literature on visual analogy problems and strategy use. This study extends early findings on differences between more and less successful problem solvers to demonstrate that one factor influencing the selection of a particular solution strategy is the WMC of the individual.

It is worth noting that this study did involve creating a new version of RAPM problems and differences in overall levels of performance from the original RAPM were found. This was not completely unexpected given previous evidence that item difficulty can be related to the number of potential responses in an item (Bethel-Fox et al., 1984). As such, lowering the number of potential responses from 8 to 4 may be responsible for better performance on the new RAPM items. However, while the items may have become easier, it has been shown on numerous occasions that an RAPM item's difficulty is not related to its ability to predict WMC (Salthouse, 1993; Unsworth & Engle, 2005; Wiley et al., 2011). Also, items in the high salience condition remained strong predictors of an additional test of gF maintained.

Although these experiments clearly demonstrate that susceptibility to distraction is affecting RAPM problem solving, which is consistent with an attentional control explanation for the relation between WMC and RAPM, it is important to bear in mind that attentional control is probably not the only factor underlying the WMC–RAPM relationship. For example, Unsworth and Spillers (2010) have demonstrated that even after controlling for attentional control, measures of WMC still explain substantial variance in a gF construct. In fact, WMC measures continue to explain the gF construct after additionally accounting for variance related to efficient retrieval from secondary memory. Similarly, work by Cowan et al. (2005) has demonstrated a relationship between the scope of attention (i.e., the immediate capacity of attention) and performance on the Raven Progressive Matrices. While the current work has demonstrated that susceptibility to distraction explains some of the variance between WMC and RAPM, it is likely only one of a constellation of factors needed to fully explain the relationship.

In summary, the findings above make several notable contributions to the literature. First, these results demonstrate that susceptibility to distraction is partially responsible for the established relationship between performance on complex span tasks and the RAPM. Second, the results demonstrate that WMC influences solution strategies used for the RAPM. Additionally, these strategy differences may account for part of the role that the salient distractor plays in the relationship between WMC and the RAPM. These results can be well explained by an attentional control account, coinciding with research suggesting that attentional control is both represented by WMC (Engle, 2002; Kane & Engle, 2003; Kane et al., 2001) and plays an important role in its relation to gF (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Kane et al., 2004; Marshalek et al., 1983). Finally, it is also interesting to note that performance on high salience items remained a significant predictor of performance on the paper folding task after accounting for visual–spatial ability, over and above low salience item performance. This suggests that the role of attentional control in the face of distraction may go beyond the WMC–RAPM relationship, and relate to a more fundamental aspect of gF. While further research regarding the relations between measures of executive functioning and intelligence is clearly warranted, this study provides an important step in the process of understanding the relationship between executive functioning as measured by WMC and gF as measured by the RAPM.

Acknowledgments

This project was completed as part of the master's thesis of the first author. The authors thank Jim Pellegrino and Ben Storm for their feedback on this project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.intell.2012.06.001.

References


