A Mechanistic Account of the Relation between Working Memory Capacity and Fluid Intelligence

by

Kimberly Wingert

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2018 by the
Graduate Supervisory Committee:

Gene Brewer, Co-Chair
Danielle McNamara, Co-Chair
Samuel McClure
Thomas Redick

ARIZONA STATE UNIVERSITY

AUGUST 2018

ABSTRACT

Working memory capacity and fluid intelligence are important predictors of performance in educational settings. Thus, understanding the processes underlying the relation between working memory capacity and fluid intelligence is important. Three large scale individual differences experiments were conducted to determine the mechanisms underlying the relation between working memory capacity and fluid intelligence. Experiments 1 and 2 were designed to assess whether individual differences in strategic behavior contribute to the variance shared between working memory capacity and fluid intelligence. In Experiment 3, competing theories for describing the underlying processes (cognitive vs. strategy) were evaluated in a comprehensive examination of potential underlying mechanisms. These data help inform existing theories about the mechanisms underlying the relation between WMC and gF. However, these data also indicate that the current theoretical model of the shared variance between WMC and gF would need to be revised to account for the data in Experiment 3. Possible sources of misfit are considered in the discussion along with a consideration of the theoretical implications of observing those relations in the Experiment 3 data.

ACKNOWLEDGMENTS

I would like to extend a special acknowledgement to Gene Brewer, Danielle McNamara, Samuel McClure, and Thomas Redick for their helpful feedback on previous drafts of this manuscript. Parts of this research were supported by National Science Foundation grant 1632327.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Fluid intelligence (gF) has been defined as the ability to reason and solve problems in novel situations requiring the learning of complex relations (Cattell, 1971; Unsworth, Fukuda, Awh, & Vogel, 2014). As noted in Unsworth et al. (2014), gF is important because it predicts performance in educational settings (e.g., Deary et al., 2007). Given the importance of educational success, many researchers have devoted much time and resources to understanding the nature of gF and why it is related to other important performance indices. Working memory capacity (WMC) is one example of a cognitive construct that researchers have consistently shown correlates highly with gF (for a recent summary of these studies and an explanation for known variability in the strength of the relation across different studies, see Chuderski, 2013; also see Unsworth and Redick, 2017). Working memory is a system that allows for the dynamic maintenance of information in primary memory in the face of a variety of sources of distraction (Unsworth and Redick, 2017). Variability in the functioning of this system is represented by measures of WMC, which research has shown similarly relates to skills important in educational contexts such as reading comprehension (Daneman & Carpenter, 1980; Turner & Engle, 1989; but see McNamara & O'Reilly, 2009). Different theories have been proposed regarding the nature of this shared variance between gF and WMC. The goal of the present study is to evaluate these theories.

CHAPTER 2

INTELLIGENCE AND MEASUREMENT

In 1904, Charles Spearman used the term ‘g’ to describe the shared variance represented by positive correlations across a diverse array of cognitive tasks (also known as the positive manifold). According to Cattell (1971), g is the psychometric term for the factor approximating the measurement of general intelligence. However, different researchers may refer to g or intelligence interchangeably. As noted by Cattell (1971), variation in g was initially represented by variability in performance on a diverse range of tasks representing a number of skills (e.g., see *Figure* 1 for an example with tasks measuring vocabulary learning, analogies, math proficiency, and mechanical knowledge) important for educational, occupational, and life success.

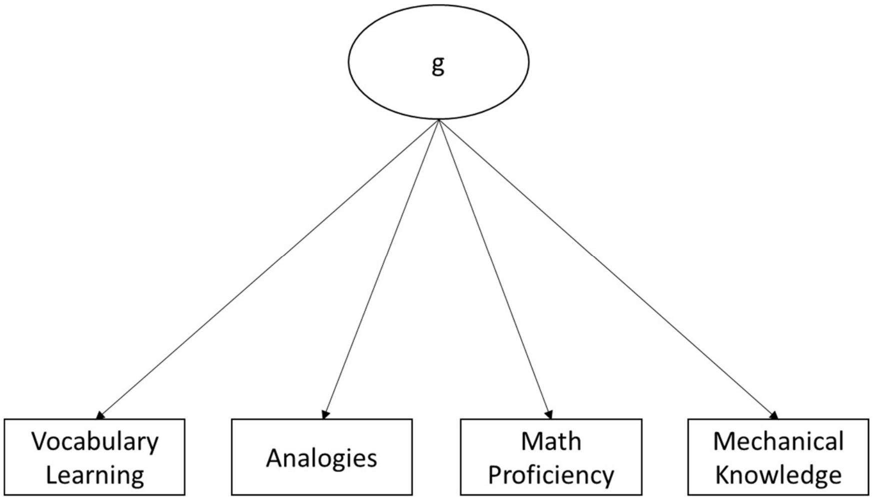


Figure 1. Example factor structure of g. g is represented by the shared variance present among indices of various skills assessed by performance on a range of tasks. The shared variance represented by g manifests as positive correlations across all tasks, which collectively is known as the positive manifold.

However, performance on these tasks is dependent upon access to learning opportunities which spurred a reconsideration of whether the existing measures underlying the positive manifold were representative of intelligence. The issue arose because individuals without such access to educational opportunities were being evaluated using these measures but then the evaluation was interpreted as if g represented something innate and unchangeable (Cattell, 1971). This interpretation is hardly fair if those without access to education would have been evaluated differently given the same opportunities as the privileged participants. As Cattell (1971) noted, this concern led to the development of perceptual or culture-free measures of intelligence.

Perceptual or culture-free measures of intelligence were generated in an attempt to create alternative measures to those that contained nonrandom error due to differences in access to educational opportunities. These alternative tasks were generated in an effort to measure fluency in educing relations in novel situations that do not rely on knowledge/skills that were acquired more readily for individuals who were afforded access to quality education. Little did anyone know that when individuals performed these tasks along with the old ones, conclusions drawn about their data would call for a shift in how researchers such as Cattell (1971) thought about g/intelligence. Cattell (1971) noted that across many studies using all of these tasks there was a clear trend in the data suggestive of two factors, fluid intelligence or gF, and crystalized intelligence or gC. That these two factors were correlated further indicated that these two factors were perhaps collectively what Spearman was intending to measure when he wrote of his theory of g (1904; see *Figure* 2). Cattell (1971) further argued that the two factors were correlated because fluid intelligence is needed to form crystallized intelligence over time with experience in learning environments such as school. According to Cattell (1971), this baseline ability to reason in new contexts thus represents a way to compare how well people will do in a variety of situations given equal opportunities to learn as others had. The new gF factor of intelligence thus appears to be the more appropriate factor of intelligence to use to evaluate individuals based on inherent unchangeable ability.

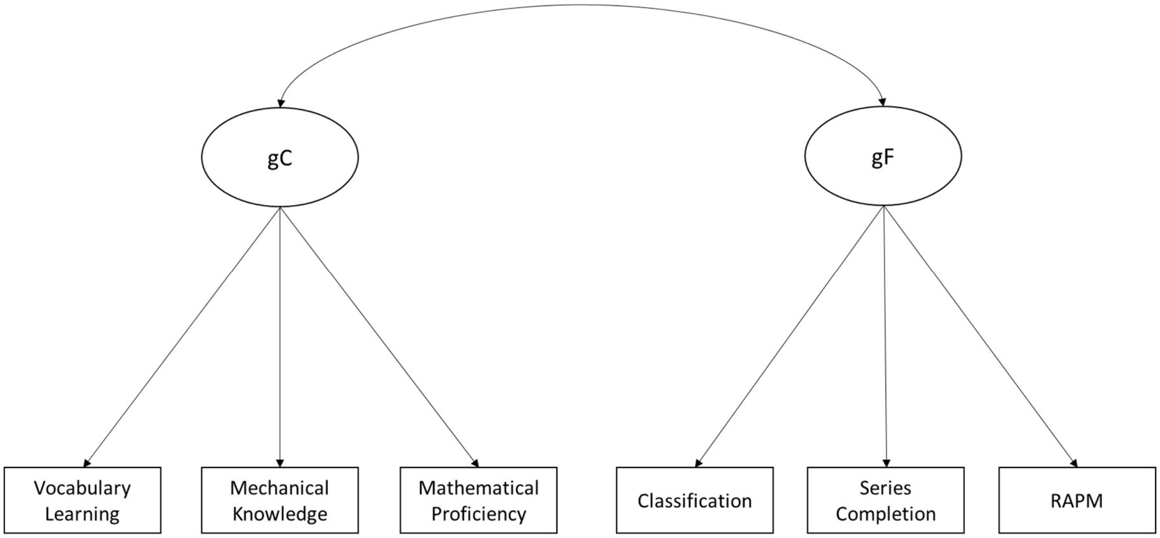


Figure 2. Example of a two-factor model of intelligence. Traditional measures of g now form a crystalized intelligence factor (gC), and the new perceptual or culture-free measures form a fluid intelligence factor (gF). The positive manifold arises because these two factors are positively correlated.

The Advanced Progressive Matrices task (RAPM; and the Progressive Matrices task) is a measure of gF that has been used frequently by researchers (e.g., see Ackerman, Beier, and Boyle, 2005 for an indication of how frequently this task is used compared to others in studies of the relation between gF and a particular measure of memory that will be discussed later). The RAPM was created by Raven (1936; 1941; 2000; Raven, Raven, & Court, 1998) to measure how well individuals are able to learn complex relations in materials they are unlikely to have previously been exposed to (the Advanced version was developed for higher ability populations such as the undergraduate students who participated in some of the research using this task). In the RAPM task, participants are asked 36 questions of increasing difficulty. On each question, a 3x3 display of geometric patterns is presented, with the bottom right pattern missing. The participant is then asked to choose the pattern that completes the overall 3x3 display from eight possible solutions. To arrive at the correct solution, participants must consider the rules that each column and each row follow and select the unique solution that completes the array without breaking a rule. The rules in RAPM are not necessarily the same across both rows and columns, and more difficult problems contain more complex rules and relations that must be considered in order to arrive at the correct solution. Accuracy on this task is simply the proportion of problems the participant answers correctly out of 36 total problems. Other tasks measuring fluid intelligence similarly require the participant to reason in novel situations, with novelty representing a critical condition thought to remove the contribution of unequal access to educational opportunities to intelligence scores.

CHAPTER 3

INDIVIDUAL DIFFERENCES IN INTELLIGENCE AND WMC

Cattell (1971) proposed that gF is needed to form gC over time, leading to the correlation between the factors which results in the positive manifold. Thorsen, Gustafsson, and Cliffordson (2014) provided support for Cattell's (1971) theory that gF is needed to form gC over time by demonstrating that gF has a continuous influence on gC for students evaluated at grades three, six, and nine. Cattell (1971) further noted that measures of intelligence were often used by employers and schools to evaluate the potential of an individual for success in their organization. Given the immense amount of time that it would take to have each potential employee or student complete all of the tasks that correlate in the positive manifold, researchers shifted focus to the finding the underlying sources of the positive correlations found across performance on all intelligence tasks so that the shared variance representing intelligence can be measured with more reliable, valid, and with fewer tasks.

An approach to theory construction advocated by Underwood (1975) involves generating theories that make specific predictions about how two constructs should be related. Individual differences in one construct should be related to individual differences in the other construct to the degree that the two constructs are thought to rely on similar cognitive processes (also see Turner and Engle, 1989). Kane et al. (2004) demonstrated that working memory capacity was highly related to gF and was also related to gC (although WMC was not as strongly related to gC compared to gF). Thus, there is a strong overlap in cognitive processes supporting WMC and gF, and a weaker overlap in cognitive processes underlying WMC and gC. By examining why WMC is related to gF and gC researchers can begin to determine sources of variability underlying each factor of intelligence (e.g., Turner & Engle, 1989). The larger overlap in processes underlying the relation between WMC and gF (compared to gC) along with the notion that gF contributes to the development of gC (Cattell, 1971; Thorsen et al., 2014) led to a heightened interest in determining why WMC is related to gF.

The relation between WMC and gF has ranged from modest (Ackerman et al., 2005) to strong (Kane, Hambrick, & Conway, 2005) in the literature (also see Unsworth and Redick, 2017). As a result, multiple lines of research have examined the nature of the processes underlying this shared variance. WMC estimates describe variability in the functioning of the working memory system and thus represent individual differences in how well the system is able to manage the contents of primary memory in distracting environments. Unsworth (2016) proposed a multifaceted view of WMC with variance in WMC arising due to variation in primary memory capacity, attention control, and cue-dependent retrieval of task-relevant information from secondary memory. Unsworth, Fukuda, Awh, and Vogel (2014) further demonstrated that the shared variance between WMC and gF arises due to variation in all three of these processes (see *Figure* 3). These three cognitive processes are distinct, important for regulating WMC, and represent important sources of variability in gF.

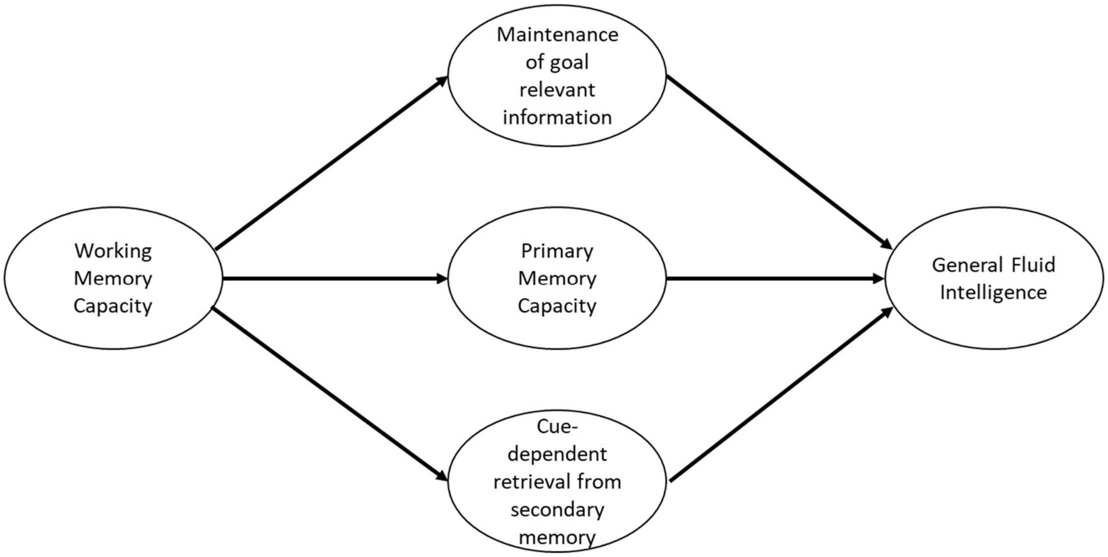


Figure 3. Model of the shared variance between WMC and gF. Working memory capacity and gF are related because both rely on the maintenance of goal relevant information, primary memory capacity, and the cue-dependent retrieval of momentarily displaced information from secondary memory (e.g., see Unsworth et al., 2014).

The overlap in the processes required to complete working memory tasks (such as complex span tasks) and gF tasks (such as the RAPM) highlights the value in understanding the multifaceted nature of complex span tasks. Complex span tasks intersperse to-be-remembered information with a processing task which is meant to be distracting. These tasks were designed to more closely approximate the active processing needed to distinguish working memory from traditional definitions of short-term or primary memory that were defined by their storage requirements in the absence of online processing (e.g., Baddeley & Hitch, 1974; also see Unsworth & Engle, 2007a and Unsworth & Redick, 2017). For example, the reading span task (see Daneman & Carpenter, 1980; Turner & Engle, 1989) intersperses sentences that a participant must decide are valid or invalid with letters that a participant must later recall in the correct serial order. Thus, a participant may see a sentence such as “I like to run in the park” and must indicate via a mouse click whether the sentence is TRUE (makes sense) or FALSE (does not make sense; in this example the correct answer is TRUE).

After the participant indicates whether a sentence is valid, they are presented with a letter for a fixed amount of time (which is 1 second for the letters in this particular complex span task), followed by another sentence that a participant must judge. A trial consists of sets of sentences and letters that alternate, and the number of sets varies randomly from three to seven with the constraint that the task contains three presentations of each list length (size of the set). At the end of a trial, participants are asked to recall the letters presented in serial order, and the number of letters that a participant recalls in the correct serial position is taken as an estimate of the individual’s working memory capacity (partial-unit span scoring; see Conway et al., 2005). Although the nature of the processing and storage tasks changes with variants of complex span tasks, the critical requirement to maintain access to task-relevant information in the face of ongoing processing of distracting information is retained.

As summarized by Engle and Kane (2004), there are at least two approaches to describing why complex span tasks relate to gF (also see Unsworth & Redick, 2017). According to the first approach, individual differences in working memory capacity relate to broader cognitive ability due to the interaction between attention and memory processes (e.g., Unsworth et al., 2014). This approach critically highlights the importance of both the processing and storage components of working memory tasks (e.g., Baddeley & Hitch, 1974; Daneman & Carpenter, 1980). Researchers such as McNamara and Scott (2001) provided evidence that participants can improve the storage capacity of the working memory system by learning effective strategies for maintaining information in memory (e.g., elaboration rather than rote rehearsal). However, as noted by Engle and Kane (2004), the relation between WMC and performance on higher-order cognitive tasks actually increases when differences in strategies are controlled for, indicating that differential use of strategies is a nuisance variable that reduces the shared variance between the tasks. Thus, it appeared that strategy use was not an important factor in the shared variance between WMC and gF. It was a little over a decade before an argument was presented that attempted to invalidate this conclusion drawn by Engle and Kane (2004).

In response to the argument put forth by Engle and Kane (2004), Gonthier and Thomassin (2015) pointed out a flaw in previous studies examining the contribution of differences in strategic behavior to the relation between WMC and gF. Specifically, these studies measured strategy use in working memory tasks, which should relate to performance on working memory tasks but should not facilitate performance on gF tasks that do not usually benefit most from the direct application of strategies that improve WMC. Rather, as Gonthier and Thomassin (2015) argued, it is the strategic behavior adopted to complete the gF task rather than the strategic behavior adopted to complete the working memory tasks that should be measured in order to critically evaluate the claim that WMC predicts performance on gF tasks due to differences in strategy use. This critical argument arises because the theoretical stance being evaluated is that variability in the functioning of the working memory system leads to differences in strategic behavior which determine performance on gF tasks like the RAPM.

If the strategies used to complete a working memory task were the exact same as the strategies used to complete the RAPM task, the tasks would be isomorphic. However, Ackerman et al. (2005) and Chuderski (2013) systematically examined evidence that the tasks measure the same thing and provided evidence refuting the claim that WMC and RAPM tasks are isomorphic. While the strategies may share variance, the strategies used to complete an RAPM task differ from those used to complete a working memory task. While working memory capacity is improved when participants are taught to use more effective strategies during a memory task such as using elaboration rather than rote rehearsal (McNamara and Scott, 2001), learning to use strategies such as elaboration to facilitate recall is not necessarily going to facilitate the ability to reason in novel situations. In fact, Bethell-Fox, Lohman, and Snow (1984; also see Vigneau, Caissie, & Bors, 2006) discuss two different strategies adopted by individuals to complete a gF task that do not clearly map on to how one would naturally perform a complex-span task.

Higher ability participants tend to use a constructive matching strategy to complete the RAPM task (Vigneau et al., 2006). The constructive matching strategy involves generating the correct solution to the RAPM problem prior to examining the response alternatives. Once a potential solution has been decided, the possible solutions can be examined for the constructed solution that was generated by the individual. The second strategy used by lower ability participants involves examining the response alternatives one by one, ruling those out that do not appear to match and selecting the solution that appears to follow the rules of the RAPM problem (Bethell-Fox et al., 1984; Vigneau et al., 2006). If individual differences in WMC lead to differences in strategic behavior on a RAPM task and these differences in strategic behavior are why WMC is related to performance on the RAPM, then individual differences in WMC should predict whether participants adopt a constructive matching or response elimination strategy and the relation between WMC and RAPM task performance should disappear or decrease if differences in the use of these strategies is controlled (Gonthier & Thomassin, 2015, following the same logic presented in Engle & Kane, 2004 applied to more appropriate measures; however, see Rucker, Preacher, Tormala, & Petty, 2011).

Initial evidence that individual differences in WMC are related to the type of strategy adopted to complete an RAPM task was presented by Jarosz and Wiley (2012), who found support for the idea that high WMC participants are more likely to adopt a constructive matching strategy and low WMC participants are more likely to adopt a response elimination strategy by examining eye tracking data collected while participants completed the RAPM task. Following up on this work, Gonthier and Thomassin (2015) conducted two studies in order to examine whether the shared variance between WMC and RAPM task performance arises due to differential use of these two strategies or if strategic behavior represents little more than a nuisance variable (i.e., Engle & Kane, 2004; also see Unsworth & Redick, 2017). The two studies conducted by Gonthier and Thomassin (2015) heeded the call to unite correlational and experimental methods to provide evidence that two constructs are in fact related and that the reason for the relation is known (Cronbach, 1957).

If differences in strategic behavior are truly the reason for the relation between WMC and RAPM task performance, then if all participants are taught the more effective constructive matching strategy and the structure of the task promoted the use of this strategy, the relation observed by many researchers (e.g., see Ackerman et al., 2005 and

Kane et al., 2005) should be reduced or disappear. If the conclusions drawn by Engle and Kane (2004) that strategic behavior represents a nuisance variable are correct, then the correlation between WMC and RAPM task performance should increase. In the first experiment conducted by Gonthier and Thomassin (2015), participants were instructed to use a constructive matching strategy in one condition and were not given constructive matching instructions in a control condition. Additionally, in the constructive matching condition the possible solutions to each RAPM problem did not appear until 15 seconds after the problem appeared on the screen to encourage all participants to generate a solution prior to viewing the responses. In the control condition, the possible solutions appeared at the same time as the problem and participants were free to select their strategy naturally. In line with the hypotheses of Gonthier and Thomassin (2015), the correlation between WMC and RAPM task performance was lower in the constructive matching condition. Contrary to the conclusions drawn by Engle and Kane (2004), the results of Gonthier and Thomassin (2015) support the idea that differences in strategic behavior underlie the relation between WMC and RAPM task performance (but see Unsworth & Redick, 2017 and Loesche, Wiley, & Hasselhorn, 2015).

In their second experiment, Gonthier and Thomassin (2015) measured working memory capacity, strategic behavior, and gF (RAPM task performance), and their proposed model appears in *Figure* 4. As shown in *Figure* 4, the solid arrows going from WMC to constructive matching and from WMC to response elimination reflect the prediction of a significant regression path and are in line with the predictions that follow from Jarosz and Wiley (2012). More specifically, Gonthier and Thomassin (2015) predicted that WMC would be positively related to the use of a constructive matching strategy and negatively related to the use of a response elimination strategy (e.g., Jarosz & Wiley, 2012). Additionally, the solid arrows going from constructive matching to RAPM task performance, and from response elimination to RAPM task performance reflect the prediction of a significant regression path and are in line with the predictions that follow from Bethell-Fox et al. (1984) and Vigneau et al. (2006).

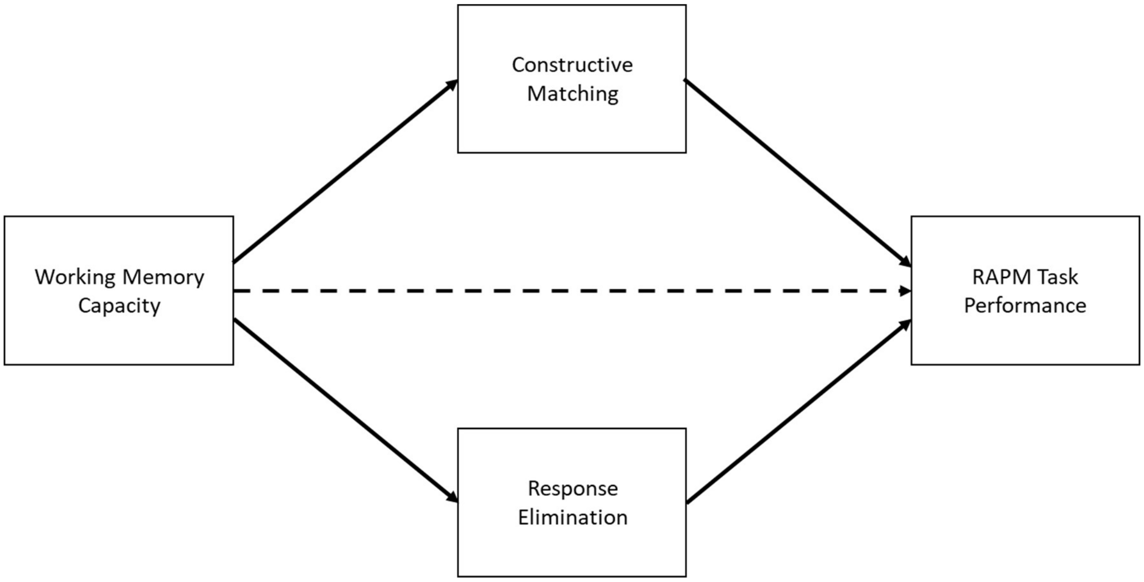


Figure 4. Model fit to the data in Gonthier and Thomassin (2015). Solid arrows correspond to a prediction of a significant regression path, whereas dotted arrows correspond to a prediction of a nonsignificant regression path. Described in detail in the text.

Gonthier and Thomassin (2015) further predicted that use of a constructive matching strategy would be positively related to RAPM task performance and use of a response elimination strategy would be negatively related to RAPM task performance (e.g., Bethell-Fox et al., 1984 and Vigneau et al., 2006). They predicted that the correlation between WMC and RAPM task performance would be significantly lower after accounting for differences in strategic behavior in this way. In fact, they found the strongest evidence for their hypothesis because individual differences in strategic behavior fully accounted for the relation between WMC and RAPM task performance (but see Rucker et al., 2011 for further consideration of the appropriateness of this approach). The relation between WMC and RAPM task performance was not present after accounting for differences in strategic behavior (the dotted line going from WMC to RAPM task performance represents a nonsignificant path).

Taken together, these data do not present evidence refuting the underlying processes that Unsworth et al. (2014) proposed account for the relation between WMC and RAPM task performance (see *Figure* 3 and Rucker et al., 2011). Rather, Gonthier and Thomassin (2015) highlight the likelihood of the alternative view that individual differences in the maintenance of information in primary memory, the capacity of primary memory, and the cue-dependent retrieval of information from secondary memory all help determine which strategy is ultimately used during the RAPM task (Unsworth et al., 2014 additionally note that other processes may covary with the cognitive processes and help explain the relation between WMC and gF). A study that includes measures of all of these underlying processes is needed to assess the levels of processing theory outlined by Gonthier and Thomassin (2015) regarding the contribution of the cognitive processes proposed by Unsworth et al. (2014) to the strategic behavior of a participant.

If the relation between WMC and RAPM task performance can no longer be explained by differences in strategic behavior after accounting for individual differences in primary memory capacity, maintenance of information in primary memory, and cue-dependent retrieval of information from secondary memory, then variation in the latter three processes are the reason for differences in strategic behavior. However, if accounting for variance in those processes only reduces the contribution of variation in strategic behavior then the other three processes present only a partial explanation for strategic differences that contribute to the relation between WMC and RAPM task performance (but see Rucker et al., 2011 for a discussion of how different amounts of measurement precision may complicate interpretation of these effects). To contradict the claims in Unsworth et al. (2014), the opposite pattern would need to be observed such that individual differences in the three processes that account for the relation between WMC and RAPM task performance in Unsworth et al. (2014) no longer account for this relation after controlling for variation in strategic behavior. Thus, the relative contribution of each underlying process can be inferred if all five of them are measured as shown in *Figure* 5.

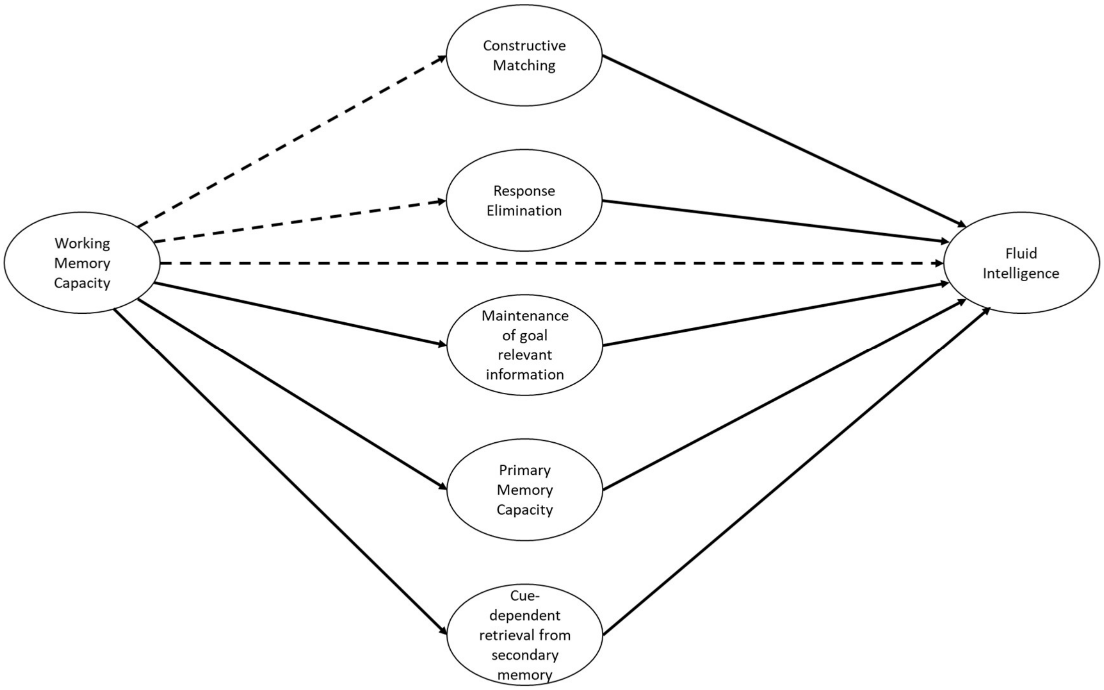


Figure 5. A model including all three cognitive processes proposed by Unsworth et al. (2014), and the two processes reflecting strategic behavior (Gonthier & Thomassin, 2015) as mediators of the relation between WMC and gF. Solid arrows correspond to a prediction of a significant regression path, whereas dotted arrows correspond to a prediction of a nonsignificant regression path. Described in detail in the text.

CHAPTER 4

THE PRESENT STUDY

The aim of the present study is to conduct three large-scale experiments to evaluate the theories outlined in this paper describing the source of the shared variance between WMC and gF (i.e., cognitive processes and strategic behavior in an RAPM task).

Experiment 1

A study conducted by Jastrzębski, Ciechanowska, and Chuderski (2017) failed to replicate the mediation results in Gonthier and Thomassin (2015), though the RAPM tasks across the two studies differed in the time that participants were allowed to complete the RAPM task (a time limit was imposed in Jastrzębski et al., 2017 but not in Gonthier & Thomassin, 2015). The goal of Experiment 1 was to replicate the results presented in Gonthier and Thomassin (2015) showing that the relation between WMC and RAPM task performance was fully explained (fully mediated) by individual differences in strategic behavior (see *Figure* 4; also see Rucker et al., 2011), and that the correlation between WMC and RAPM task performance is lower when participants learn to use a constructive matching strategy and are provided with environmental support to implement the strategy. Experiment 1 manipulated strategic behavior to determine whether there is support for the notion that the reason that WMC is related to RAPM task performance is due to differences in the use of more or less effective strategies. In a control condition the participant was allowed to naturally select the strategy used to complete the RAPM task, and in a constructive matching condition participants learned to use the constructive matching strategy. The RAPM task was also edited in the constructive matching condition such that the problem remained on the screen for 15 seconds before the possible solutions appeared (as in Gonthier and Thomassin, 2015). Two primary hypotheses and a set of related predictions follow from the goal to replicate Gonthier and Thomassin (2015).

Hypothesis 1 (H1): Retrospective reports of strategic behavior will fully mediate the relation between WMC and RAPM task performance.

As in Gonthier and Thomassin (2015), H1 can be decomposed into the following set of predictions: 1) WMC will be positively related to reported use of the constructive matching strategy, 2) WMC will be negatively related to reported use of the response elimination strategy (both 1 and 2 follow from Jarosz and Wiley, 2012), 3) reported use of the constructive matching strategy will be positively related to RAPM task performance, 4) reported use of the response elimination strategy will be negatively related to RAPM task performance (both 3 and 4 follow from Bethell-Fox et al., 1984 and Vigneau et al., 2006), 5) the indirect path (mediation) from WMC to RAPM task performance via both reported strategic behaviors will be significant, and 6) the correlation between WMC and RAPM task performance will no longer be significant after accounting for differences in strategic behavior (see *Figure* 4; also see Rucker et al., 2011 for a discussion of the appropriateness of this final prediction).

Hypothesis 2 (H2): The relation between WMC and RAPM task performance will decrease in the constructive matching condition.

Gonthier and Thomassin (2015) predicted that the relation between WMC and RAPM task performance can be explained by individual differences in strategic behavior. However, the correlation between WMC and RAPM task performance was still significant in the constructive matching condition in their study. This could reflect contributions of other processes to the shared variance or could reflect the persistent use of response elimination strategies even when they are not shown the possible solutions immediately. Thus, H2 can be decomposed into four predictions: 1) the correlation between WMC and RAPM task performance will be lower in the constructive matching condition as in Gonthier and Thomassin (2015), or 2) the correlation between WMC and RAPM task performance will not be significant in the constructive matching condition. In both cases the theory put forth by Gonthier and Thomassin (2015) would be supported. However, if 3) the correlation between WMC and RAPM task performance is unchanged, or 4) increases, this would support the idea that individual differences in strategic behavior are a nuisance variable rather than an underlying mechanism (Engle & Kane, 2004). However, an alternative explanation for any inconsistencies across this study and Gonthier and Thomassin (2015) is that retrospective reports of strategic behavior are given at the end of the RAPM task and strategic behavior may not be consistent across problems within a participant.

Experiment 2

Bethell-Fox et al. (1984) indicated that a response elimination strategy was used by low ability participants on more difficult trials (also see Vigneau et al., 2006). Thus, a better understanding of the contribution of strategic behavior to accounting for the relation between WMC and RAPM task performance could be gained by asking participants to report the strategy they used after every RAPM problem. The goal of Experiment 2 is to determine how intraindividual differences in strategic behavior across RAPM problems of increasing difficulty varies as a function of strategy condition

(Strategy: Constructive Matching vs. Control). If intraindividual differences in strategic behavior vary across conditions, then characteristics of the sample that differ from the sample in Gonthier and Thomassin (2015) may lead to a failure to replicate their significant mediation effects. By examining strategic behavior on each trial rather than at the end of the experiment, the measure of strategic behavior can more accurately reflect the behavior endorsed during the task.

Figure 6 illustrates the idea that response elimination is used on more difficult trials by low ability individuals (Bethell-Fox et al., 1984). The point at which a low ability individual switches to the response elimination strategy may thus be a better index of differences in strategic behavior because if the data appear as in *Figure* 6 there is still overlap in the use of the constructive matching strategy across high and low ability individuals on easier problems. The overlap in the use of the constructive matching strategy on easier problems introduces unexplained variance to data that were intended to assess whether WMC and RAPM performance are related due to individual differences in strategic behavior. A measure of differences in strategic behavior should be of a sufficiently narrow scope to remove what is common to gain a less noisy measure of differences in strategic behavior.

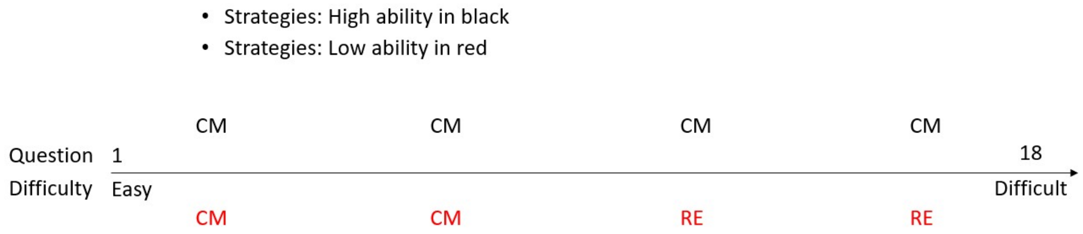


Figure 6. RE is used by low ability participants on more difficult trials. Described in detail in the text.

Vigneau et al. (2006) argued that the high reliability of their measures of strategic behavior across RAPM problems of increasing difficulty is inconsistent with the notion of a switch point._Additionally, although the switch point does not contain nonrandom error due to similarity in the type of strategy used across participants on easier problems in the RAPM task, it may contain other types of error and a switch point ignores one of the benefits of examining the strategy used on each RAPM problem such as the ability to estimate variability within a participant as a measure of strategic behavior. That is, a low ability participant may switch to a response elimination strategy on a specific trial as shown in *Figure* 6 (Bethell-Fox et al., 1984), but on the next problem they may be able to implement the constructive matching strategy again. Jastrzębski et al. (2017) indicated that it is possible that WMC-related differences in mental resources may lead to differences in how consistently a constructive matching strategy can be implemented. The model in *Figure* 4 representing the critical predicted mediation effect is updated in *Figure* 7 to reflect the use of new measures of differences in strategic behavior that represent intraindividual variability in strategic behavior across RAPM problems of increasing difficulty.

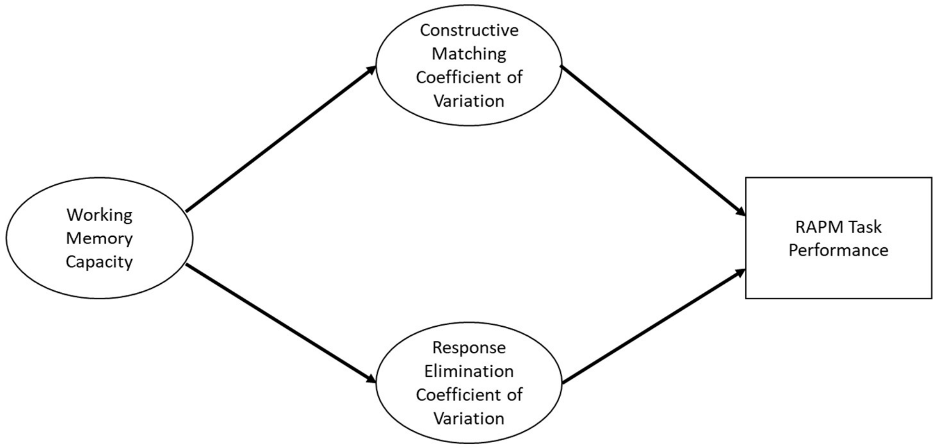


Figure 7. A revised model of the shared variance between WMC and RAPM task performance with differences in strategic behavior that should increase the reliability and validity of the measurement of these differences. Described in detail in the text.

Similar to the dotted arrow in *Figure* 4 going from WMC to RAPM task performance reflecting a nonsignificant relation in the data presented in Gonthier and Thomassin (2015), the absence of a path going from WMC to RAPM task performance in *Figure* 7 indicates the a priori prediction that the path will not be significant after accounting for individual differences in strategic behavior. As in Experiment 1, there are two primary hypotheses that follow from the goal to assess the hypotheses put forth in Gonthier and Thomassin (2015) using measures that may be more appropriate for examining differences in strategic behavior.

Hypothesis 3 (H3): Intraindividual variability in strategic behavior will fully mediate the relation between WMC and RAPM task performance.

Similar to the predictions following H1, H3 can be decomposed into the following set of predictions: 1) WMC will be negatively related to variability in the application of constructive matching and response elimination strategic behavior, 2) variability in constructive matching strategic behavior and variability in response elimination strategic behavior will be negatively related to performance on the RAPM task, 3) the indirect paths from WMC to RAPM task performance via the two indices of variability in strategic behavior will be significant, and 4) the correlation between WMC and RAPM task performance will no longer be significant after accounting for differences in strategic behavior (see *Figure* 7; however, also see Rucker et al, 2011).

Hypothesis 4 (H4): The relation between WMC and RAPM task performance will decrease in the constructive matching condition.

Experiment 2 addresses the concern that retrospective reports of overall strategy use on the task may introduce nonrandom error that may vary across samples. That is, an absence of a significant mediation effect in Experiment 1 would not necessarily support the notion that differences in strategic behavior are a nuisance variable (e.g., Engle & Kane, 2004). The reason this conclusion would not be warranted is due to the selection of a variable representing strategic behavior that contains information from trials in which all participants may have been using similar strategies (for the easier RAPM problems), which was corrected in Experiment 2. Thus, H4 can be decomposed into four predictions (which do not differ from the predictions that follow H2): 1) the correlation between WMC and RAPM task performance will be lower in the constructive matching condition as in Gonthier and Thomassin (2015), or 2) the correlation between WMC and RAPM task performance will not be significant in the constructive matching condition. However, if 3) the correlation between WMC and RAPM task performance is unchanged, or 4) increases, this would support the idea that individual differences in strategic behavior represent a nuisance variable rather than an underlying mechanism (Engle & Kane, 2004).

Experiment 3

Thus far only performance on the RAPM task has been considered, and the cognitive processes that Unsworth et al. (2014) posited underlie the relation between WMC and RAPM task performance have not been considered alongside differences in strategic behavior. The goal of Experiment 3 is to determine the mechanisms underlying the relation between WMC and gF using a more diverse set of tasks representing gF in order to reduce the contribution of task-specific variance compared to a measure of gF solely comprised of RAPM task performance. Notably, for each of the gF tasks in Experiment 3, a set of possible solutions to a problem are available to select from, so it is reasonable to predict similar constructive matching and response elimination strategic behavior given the applicability of these strategies to solving multiple choice tests (Vigneau et al., 2006).

In Experiment 3, performance on multiple tasks measuring each construct (WMC, constructive matching strategy retrospective reports, response elimination strategy retrospective reports, primary memory capacity, maintenance of information in primary memory, cue-dependent retrieval of information from secondary memory, and gF) was assessed as a part of National Science Foundation grant 1632327[1]. The model presented in *Figure* 5 was fit to this data with each of the latent constructs represented by labeled circles in the figure representing the shared variance across three measures of that

---

[1] These measures were selected from the full set of tasks included in the NSF grant. The full set of tasks appear in *Table* 7, with the measures to be used in this study appearing in bold.

construct. Two primary hypotheses follow the suggestion in Gonthier and Thomassin (2015) that variability in WMC leads to individual differences in cognitive processes which in turn lead to the use of more or less effective strategies in a gF task.

Hypothesis 5 (H5): Cognitive processes will mediate the relation between WMC and gF

To directly assess whether the three cognitive processes outlined in Unsworth et al. (2014) contribute to the shared variance between WMC and gF, the indirect path between WMC and gF via each of these cognitive processes was examined. This hypothesis indicates that cognitive processes will mediate the relation between WMC and gF. If the indirect effects via strategic behavior are also not significant, this would provide the strongest support for the ideas proposed in Unsworth et al. (2014) and present evidence supporting the ideas discussed by Gonthier and Thomassin (2015) about why individual differences in strategic behavior arise. If the indirect effects via strategic behavior are significant in a model excluding cognitive processes but are not significant in a model including them, then this would support the idea that individual differences in cognitive processes underlying WMC drive individual differences in strategic behavior on gF tasks (however, following Rucker et al., 2011 it would also be possible to obtain this pattern of results due to differences in measurement precision across cognitive and strategic processes measures).

An alternative possibility is that cognitive processes will not mediate the relation between WMC and gF. There are two reasons that this may occur: 1) cognitive processes no longer mediate the relation between WMC and gF after controlling for individual differences in strategic behavior. This would suggest that cognitive processes are only important for describing the shared variance between WMC and gF due to their contribution to strategic behavior. This would contradict the claim made by Engle and Kane (2004) that strategies are a nuisance variable. Alternatively, 2) cognitive processes may not mediate the relation between WMC and gF even when individual differences in strategic behavior are not included in the model. This would contradict the evidence presented in many lines of work (Engle, Tuholski, Laughlin, & Conway, 1999; Unsworth & Engle, 2007b; Unsworth et al., 2014; Shipstead, Harrison, & Engle, 2015; Unsworth & Redick, 2017) suggesting that primary memory capacity, maintenance of information in primary memory, and cue-dependent retrieval of information from secondary memory all underlie the relation between WMC and gF. Collectively, these three experiments will enhance understanding of the underlying mechanisms driving the relation between WMC and gF.

Hypothesis 6 (H6): Individual differences in strategic behavior will not mediate the relation between WMC and gF after accounting for individual differences in cognitive processes

Gonthier and Thomassin (2015) highlighted the fact that their findings were not inconsistent with the theory of the relation between WMC and gF described in Unsworth et al. (2014). In particular, they summarize literature supporting the idea that RAPM task performance may be driven by individual differences in strategic behavior that arise due to individual differences in the cognitive processes underlying WMC. This leads to the prediction that 1) the regression paths going from WMC to constructive matching and response elimination strategy retrospective reports will no longer be significant after controlling for individual differences in cognitive processes. Vigneau et al. (2006) noted that it may lead to unintended consequences to assume that individual differences in strategic behavior necessarily overlap with individual differences in WMC. As a result, the model in *Figure* 5 illustrates the prediction that 2) individual differences in strategic behavior may share variance with gF independent of WMC (solid arrows going from strategic behavior to gF). These predictions are consistent with the ideas outlined in Gonthier and Thomassin (2015) and the multifaceted view of WMC proposed by Unsworth (2016), along with the view that the variance WMC shares with gF is due to variation in cognitive processes (Unsworth et al., 2014).

If individual differences in strategic behavior still mediate the relation between WMC and gF after accounting for individual differences in cognitive processes, then differences in strategic behavior are driven by some other process underlying WMC (perhaps one of the other cognitive processes discussed in Unsworth & Redick, 2017). If the mediation effect is reduced but still present, this would indicate that cognitive processes as well as some other process underlie the relation between WMC and gF. If the indirect effect is just as strong in a model that accounts for individual differences in cognitive processes, this would indicate that individual differences in strategic behavior do not arise due to individual differences in cognitive processes (inconsistent with the ideas presented in Gonthier and Thomassin, 2015). Finally, if individual differences in strategic behavior are able to explain more of the shared variance between WMC and gF after accounting for individual differences in cognitive processes, then this would indicate that cognitive processes are nuisance variables rather than important processes underlying the relation (inconsistent with Unsworth et al., 2014).

CHAPTER 5

EXPERIMENT 1 METHODS

Participants

A total of 343 participants were recruited from the introductory psychology research participation pool at Arizona State University. One participant had missing data on at least one of the tasks and was thus not included in the analyses. As noted by Draheim, Harrison, Embretson, and Engle (2017), participants that do not perform well on the processing task for the working memory measures are typically excluded because it is the processing task demands in complex span tasks that ensure these measures have construct validity as a measure of working memory rather than short-term memory. As a result, an additional 21 participants with less than 80% accuracy on the processing task were excluded. An additional 19 participants were not included in analyses because their performance on at least one of the complex span tasks was greater than 3 standard deviations from the mean. The remaining sample was examined for any remaining outliers using mahalanobis distance outlier detection, which revealed 2 additional participants whose data were also excluded. This resulted in a final sample of 300 participants included in analyses. There were 198 participants in the constructive matching condition, and 102 participants in the control condition[2].

---

[2] The sample sizes are uneven across conditions because there were originally two versions of the control condition. In the other control condition not reported in this paper, we used the traditional RAPM task similar to the control condition reported in this paper. However, we were concerned about the possibility that time spent on each problem would be confounded with our experimental manipulation. Although the data from the control condition that did not address the time confound is not presented here, it is available upon request. We were not interested in looking at the effect of the manipulation on time and thus the control condition confounded with time was dropped and the constructive matching condition data that was collected to compare to that control condition was combined with the other constructive matching data we collected to compare to the data from the control condition that controlled for time (which is presented in this paper).

Procedure

All participants consented to participate in accordance with the standards of the Arizona State University's Institutional Review Board. After consenting to participate, all participants completed the following tasks in order: 1) Operation Span, 2) Reading Span, 3) Symmetry Span, 4) Raven's Advanced Progressive Matrices (RAPM), and 5) a post-experimental questionnaire.

Materials

Operation span. The operation span complex span task is a measure of working memory capacity that intersperses to-be-remembered information with a processing task designed to prevent rehearsal of the to-be-remembered information (Turner & Engle, 1989; Unsworth, Heitz, Schrock, & Engle, 2005; Draheim et al., 2017). In operation span, the information participants need to remember are letters and the processing task consists of verifying whether a mathematical statement is true or false (5 + 2 = 2, True or False?). After practicing the task, participants first see a math operation and then are presented a letter. During a trial these math operations and letters alternate in set sizes ranging from 4 to 6. Upon completion of a trial, participants are asked to recall the letters in serial order. Each list length (4-6) was presented two times for a total possible score of 30 (Oswald, McAbee, Redick, & Hambrick, 2015). To perform this task well, participants must maintain goal-relevant information in primary memory in the face of distraction, and use cues to strategically retrieve information from secondary memory, limited by their individual primary memory capacity. A participant's score on this task was the proportion of items recalled correctly in the correct serial position (partial-unit span score).

Reading span. The reading span complex span task (Daneman & Carpenter, 1980; Unsworth, Redick, Heitz, Broadway, & Engle, 2009) was the same as the operation span task except for the processing task. During the processing task, participants were asked to determine whether a sentence made sense or not. Half of the sentences made sense, and the half that did not were created by replacing a word in a sentence that otherwise made sense with another word that rendered the sentence nonsensical. This task had the same number of trials as operation span, and the scoring procedure was the same.

Symmetry span. The symmetry span task (Shah & Miyake, 1996; Unsworth et al., 2009) was conceptually similar to operation and reading span, but the information participants were asked to remember was spatial locations presented in red in a 4x4 grid. Additionally, the processing task required participants to judge whether patterns were symmetrical around the vertical center. In this version of the task, list lengths varied from 3-5 and were presented twice each for a total possible score of 24. Similar to the other tasks, the score for each participant was calculated as the proportion of squares remembered in the correct serial position.

RAPM. In the version of the task used in Experiment 1, participants completed the 18 odd numbered items, and they had an unlimited amount of time to finish (as in Gonthier & Thomassin, 2015). Additionally, two versions of RAPM were created: the first induced all participants to use a constructive matching strategy, and the second remained the same as the original RAPM except for a couple of alterations made to ensure the tasks were comparable across conditions.

Similar to Gonthier and Thomassin (2015), to induce all participants to use a constructive matching strategy, the computerized RAPM task was revised so that each problem was displayed for 15 seconds before the 8 possible answers were displayed and a response was allowed. This manipulation was done to ensure that they considered the problem itself for a reasonable amount of time before examining the response options. To control for the possibility that people in the control condition may answer the problems much faster, a change was made to the original RAPM task such that the problem and the answers were displayed on the screen for 15 seconds before participants were allowed to answer[3]. Two other slight alterations were made to both tasks. First, the instructions were altered to be highly similar across conditions but the instructions that described constructive matching were removed for the control condition (e.g., instructions deleted asked participants to imagine the answer in their mind; see *Figure* A in the appendix). Another change made to the RAPM task signaled to the participant when they could provide their answer to the problem. During each trial (for each question), in the upper right-hand side of the screen a small red square indicated that participants could not respond yet. After 15 seconds passed, a green square appeared and participants were allowed to make a response.

Post-experimental questionnaire. Similar to Gonthier and Thomassin (2015), participants were asked two questions designed to measure response elimination, and two

---

[3] As noted in a previous footnote, we also collected data in a control condition that did not control for the 15 seconds participants had to wait in the constructive matching condition. However, our concern about needing to have a 15 second wait in both conditions appeared valid. Participants responded much more quickly when they didn't have to wait 15 seconds, and thus we chose not to include these data because any results we interpreted could be either due to a more efficient strategy in the constructive matching condition, or due to overall time spent thinking about the problem.

designed to measure constructive matching. They were further asked four filler questions and these eight questions were randomly presented for each participant. The final question was always asked last and asked the participants if the first set of questions allowed them to indicate the strategy they used. Participants answered each question by indicating how much they agreed with each statement on a scale from 1 (Not at all true) to 9 (Completely true). These questions are displayed in *Table* 1.

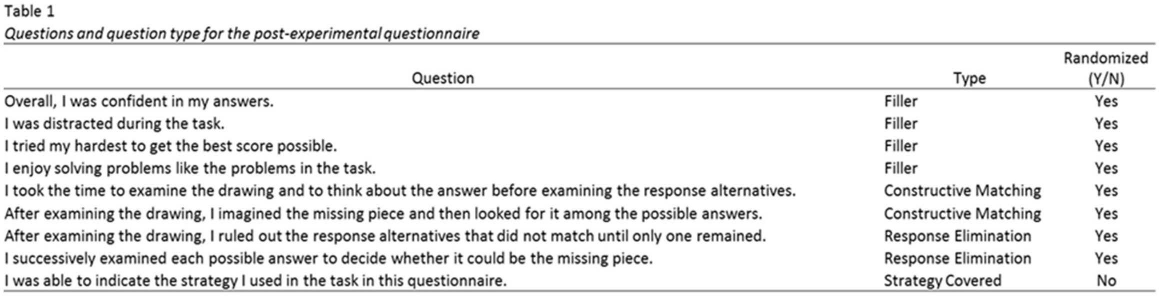


Table 1

Questions and question type for the post-experimental questionnaire

Question	Type	Randomized (Y/N)
Overall, I was confident in my answers.	Filler	Yes
I was distracted during the task.	Filler	Yes
I tried my hardest to get the best score possible.	Filler	Yes
I enjoy solving problems like the problems in the task.	Filler	Yes
I took the time to examine the drawing and to think about the answer before examining the response alternatives.	Constructive Matching	Yes
After examining the drawing, I imagined the missing piece and then looked for it among the possible answers.	Constructive Matching	Yes
After examining the drawing, I ruled out the response alternatives that did not match until only one remained.	Response Elimination	Yes
I successively examined each possible answer to decide whether it could be the missing piece.	Response Elimination	Yes
I was able to indicate the strategy I used in the task in this questionnaire.	Strategy Covered	No

CHAPTER 6

EXPERIMENT 1 RESULTS

Experiment 1 was designed to replicate the findings in Gonthier and Thomassin (2015) showing that 1), spontaneous strategic behavior in an RAPM task fully mediates the relation between WMC and RAPM accuracy (H1), and 2) when participants are provided with environmental support leading to the use of an effective strategy in a RAPM task, the correlation between WMC and RAPM accuracy is reduced (H2). To assess H2, all three proportional partial-unit span scores for the three WMC complex span tasks were submitted to factor analysis using maximum likelihood estimation (and varimax rotation) and factor scores were derived that included only the shared variance across the different WMC tasks which is thought to represent working memory capacity (e.g., Kane, Conway, Hambrick, & Engle, 2007). Descriptive statistics are presented for each of the variables in *Table* 2, and the correlations between the measures for both conditions appear in *Table* 3.

Table 2

Descriptive Statistics for the Replication of Gonthier & Thomassin (2015)

Task	Control M (SD)	Constructive Matching M (SD)
Operation Span Partial-Unit Span Score	0.80 (0.17)	0.80 (0.17)
Reading Span Partial-Unit Span Score	0.75 (0.17)	0.77 (0.16)
Symmetry Span Partial-Unit Span Score	0.68 (0.23)	0.67 (0.20)
Span Factor Score	-0.03 (0.81)	0.01 (0.83)
RAPM Accuracy	0.53 (0.20)	0.55 (0.17)
Response Elimination Q1	5.76 (2.43)	6.30 (2.23)
Response Elimination Q2	6.45 (2.16)	6.61 (2.31)
Constructive Matching Q1	7.38 (1.60)	7.71 (1.48)
Constructive Matching Q2	7.26 (1.94)	7.58 (1.66)
Strategy Coverage	6.77 (2.02)	6.99 (1.88)

Table 3

Correlations in the Constructive Matching (bottom triangle) and Control (top triangle) conditions

Variable No.	Variable	1	2	3	4	5	6	7	8	9	10
1	Operation Span Partial-Unit Span Score	-	**.34**	**.41**	**.89**	**.26**	-.02	-.02	.06	.11	.02
2	Reading Span Partial-Unit Span Score	**.51**	-	**.27**	**.67**	**.23**	-.04	-.16	.13	.09	.09
3	Symmetry Span Partial-Unit Span Score	**.34**	**.31**	-	**.64**	**.29**	**-.21**	-.14	.00	-.11	-.06
4	Span Factor Score	**.92**	**.76**	**.56**	-	**.33**	-.09	-.11	.09	.08	.03
5	RAPM Accuracy	.02	.01	**.20**	.06	-	-.07	.07	**.34**	**.23**	.15
6	Response Elimination Q1	.08	.09	.08	.11	.04	-	**.47**	**.26**	.19	**.25**
7	Response Elimination Q2	-.02	-.03	.08	-.01	.02	**.37**	-	**.28**	**.34**	**.36**
8	Constructive Matching Q1	.07	-.02	**.17**	.08	.10	**.26**	**.16**	-	**.45**	**.34**
9	Constructive Matching Q2	.12	.04	.06	.11	.04	**.30**	**.19**	**.33**	-	**.33**
10	Strategy Coverage	.03	-.02	.00	.01	.13	**.36**	**.28**	**.27**	**.32**	-

Note: Significant correlations are bolded. Colors are described in the text.

The variables in *Table* 2 were submitted to independent samples t-tests with the strategy condition manipulation as a between-subjects variable. There were no significant differences across strategy conditions in any of the measures presented in *Table* 2. However, there was a marginal effect of strategy condition on responses to the first response elimination question, $t(298) = 1.906$, $p = .058$. However, this trend reflected greater reported response elimination strategic behavior in the constructive matching condition. This may indicate that our constructive matching manipulation was ineffective. However, this trend may not represent a real effect and this measure of strategic behavior may not be sensitive to differences in strategic behavior within a person across trials. There was also a marginal effect of strategy condition on responses to the first constructive matching question, $t(298) = 1.780$, $p = .076$. This trend appeared to support greater use of a constructive matching strategy by participants in the constructive matching condition. Although this trend was in the correct direction, the measure of behavior still may not be sensitive to differences in strategic behavior within a person across trials. Thus, no conclusions may be drawn about the effectiveness of the constructive matching manipulation based on responses to the strategy questions.

Of greater concern is the fact that RAPM task performance did not significantly differ across conditions, $p = .448$. If a constructive matching strategy is used by participants who achieve better RAPM performance, then RAPM performance should be better when all participants are induced to use a constructive matching strategy if the manipulation was effective. However, Experiment 1 aimed to evaluate whether inducing all participants to use a constructive matching strategy would reduce the correlation between WMC and RAPM task performance. Although the manipulation did not lead to an overall effect on RAPM accuracy, there may have been a differential improvement in RAPM accuracy across strategy conditions as a function of WMC. H2 predicts that the relation between WMC and RAPM task performance will be reduced in the constructive matching condition, so it is the relation between these two measures that must be compared across conditions.

There are several interesting things to note about the correlations in *Table* 3. First, the working memory span scores all correlate with each other and the factor score representing the shared variance across all three of the complex span tasks (correlations in yellow). Second, the responses to the questionnaire items were all positively correlated likely reflecting their shared variance as a measure of strategic behavior on a RAPM task as well as a slight bias to respond positively across all items (correlations in orange)[4]. Third, the WMC factor score was significantly correlated with RAPM performance in the control condition ($N = 102$, $r = .33$, $p = .001$), but not in the constructive matching condition ($N = 198$, $r = .06$, $p = .374$). An examination of the bivariate correlations

---

[4] If this positive bias is a result of both strategies being used across different RAPM problems, then this positive bias should be strongest for low working memory capacity participants but weak or not present for high working memory capacity participants (who should be less likely than low working memory capacity participants to use both strategies across RAPM problems). The absence of a positive bias would result in negative correlations between the constructive matching and response elimination measures. Thus, the correlation between the constructive matching and response elimination measures should be positive for low working memory capacity participants in the control condition, and the correlation between the constructive matching and response elimination measures should be zero or negative for high working memory capacity participants in the control condition.

Participants were split into quartiles based on their WMC, and the correlations between the measures of strategic behavior (the average of the two constructive matching and the average of the two response elimination variables) were examined separately for participants in the lower quartile (low WMC; N = 25) and participants in the upper quartile (high WMC; N = 25). Average retrospective reports of constructive matching strategic behavior were positively correlated with average retrospective reports of response elimination strategic behavior for low WMC participants in the control condition, r = .60, p = .002. By contrast, average retrospective reports of constructive matching strategic behavior were unrelated to average retrospective reports of response elimination strategic behavior for high WMC participants in the control condition, r = .205, p = .325.

between the individual complex span tasks and RAPM performance revealed that it was the verbal complex span tasks specifically that lost their predictive ability once differences in strategic behavior were minimized using an experimental manipulation designed to facilitate use of the most appropriate strategy (correlations in green). Importantly, the correlation between RAPM task performance and WMC was significantly lower in the constructive matching condition, Fisher r-to-z = 2.26, *p* = .02. Thus, the data support H2 in Experiment 1, indicating that when participants are provided with environmental support leading to the use of an effective strategy in an RAPM task, the correlation between WMC and RAPM task performance is reduced (see *Figure* 8). It is worth noting that the reduction in this correlation in the constructive matching condition appeared to be driven by a selective enhancement in RAPM task performance for low working memory capacity participants.

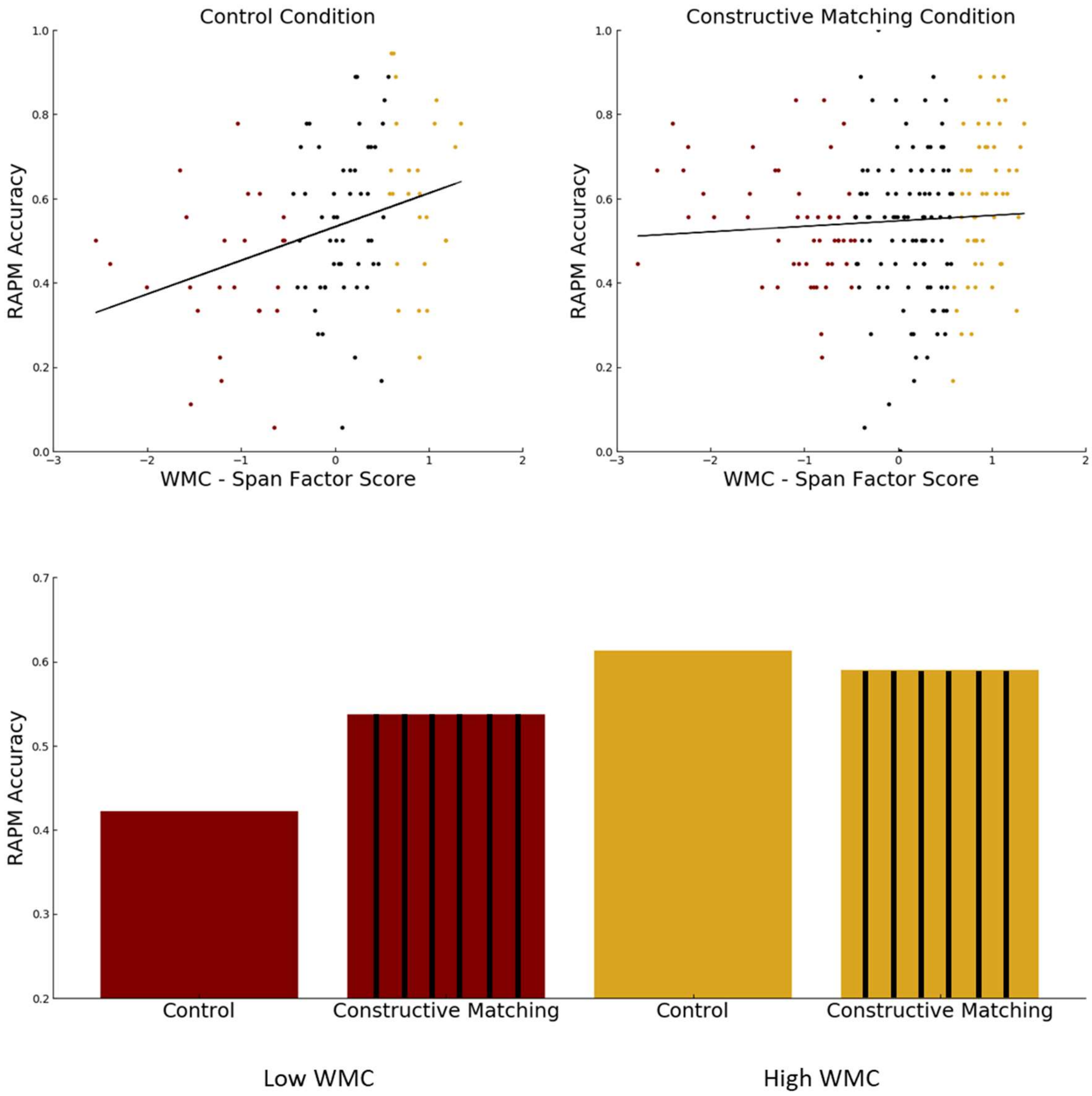


Figure 8. Top: Relation between WMC and RAPM accuracy for participants in the Control (left) and Constructive Matching (right) conditions. Bottom: Mean RAPM accuracy for low WMC and high WMC participants in the Control and Constructive Matching conditions. Described in detail in the text.

Further examination of the correlations in *Table* 3 also highlights the likelihood that the data will not support H1. The working memory capacity factor score was not correlated with any of the response elimination or constructive matching questions (correlations in blue). However, it is possible that a relation will be observed when the shared variance across constructive matching questions and the shared variance across response elimination questions is estimated in a mediation model. H1 states that individual differences in strategic behavior will fully mediate the relation between WMC and gF. To assess H1, a structural equation model was fit to the data from the control condition in MPLUS (Muthén & Muthén, 1998-2011) and appears in *Figure* 9. An aggregate strategy measure would necessarily include nonrandom error due to the use of both strategies depending on the problem. In fact, Vigneau et al. (2006) argued that aggregate measures should not be used to infer behavior because there is reason to believe that intraindividual differences underlie strategic behavior as a function of problem difficulty in a RAPM task. This bias to respond positively to using both types of strategies was accounted for in the model by allowing the constructive matching and response elimination strategic behavior latent variables to correlate for reasons outside of the model (the curved arrow in *Figure* 9).

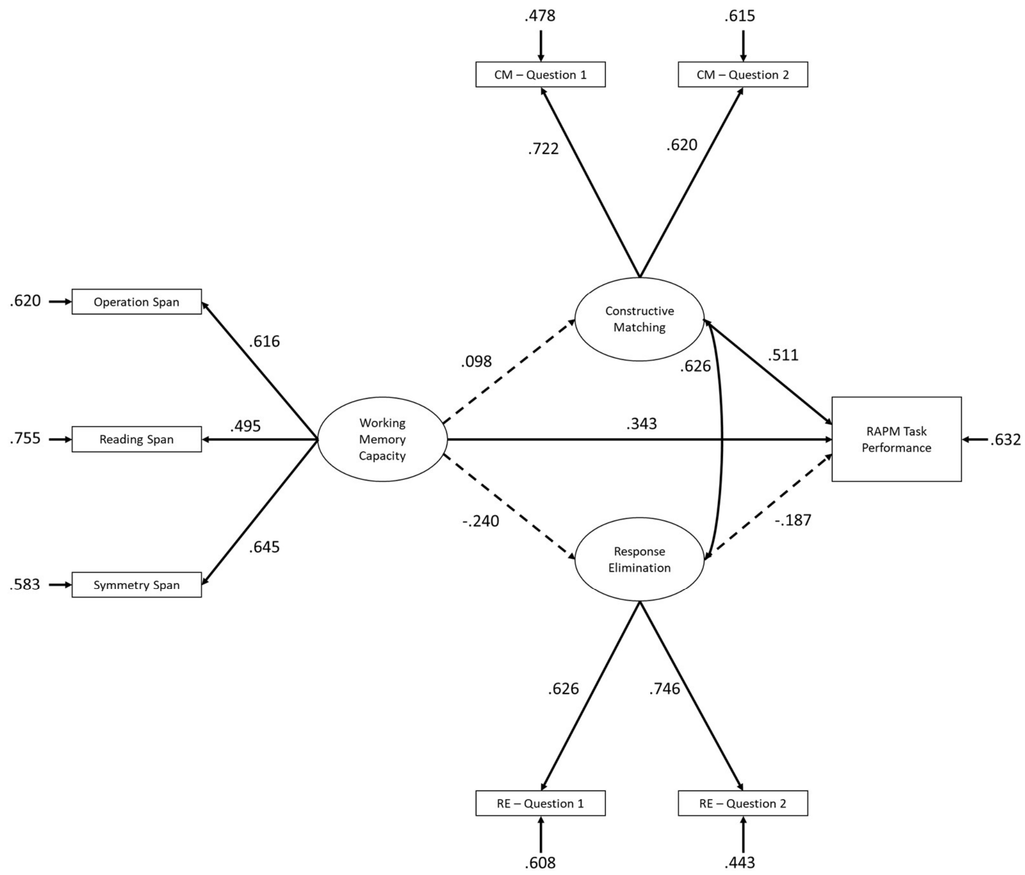


Figure 9. Structural model for Experiment 1 control condition data. Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

The model in *Figure* 9, χ^2 (15) = 16.985, p = .32, provided a good fit to the data. The model was able to recreate the sample correlation matrix over 95% of the time (RMSEA = .036 and SRMR = .047). Additionally, both the CFI (.983) and TLI (.968) indicated that the model provided > 95% improvement over a baseline model that fixes the correlations across latent factors to 0. Despite the close fit of the model to the data, the relation between WMC and RAPM task performance via constructive matching (standardized 95% bootstrapped CI[-0.262, 0.362]), and via response elimination (standardized 95% bootstrapped CI[-0.435, 0.525]) was not significant. Additionally, WMC still predicted RAPM task performance (p = .012), which is inconsistent with H1 (though see Rucker et al., 2011). However, a model omitting the direct path from WMC to RAPM task performance fit the data just as well as the model containing the direct path, $\Delta \chi^2$ (1) = 3.374, p = .07. H1 indicated that retrospective reports of strategic behavior will fully mediate the relation between WMC and RAPM task performance, as demonstrated by Gonthier and Thomassin (2015). We were unable to replicate Gonthier and Thomassin (2015), and there appeared to be a bias to respond positively on both measures of constructive matching and measures of response elimination.

Again, this supports the claim in Vigneau et al. (2006) that aggregate measures should not be used to examine when intraindividual differences in strategic behavior across RAPM problems are expected. The bias to respond positively was controlled for by allowing the strategy factors to correlate for reasons outside of the model. The amount of unexplained variance this tendency to report using both strategies added may have reduced power to detect an effect if there is one. As shown in *Figure* 9, WMC did not predict either constructive matching or response elimination. However, the direction of the regression coefficients is consistent with the predictions in the literature and the effects may not be significant because the measure has poor construct validity which reduces reliability of the effect across samples. The model in *Figure* 9 also indicates that individual differences in the use of a constructive matching strategy were positively related to RAPM task performance (p < .01). However, as shown in the figure individual differences in the use of a response elimination strategy did not predict RAPM task performance, though the regression coefficient is negative as would be predicted (Bethel-Fox et al., 1984; Vigneau et al., 2006; Gonthier & Thomassin, 2015).

The fact that response elimination strategic behavior did not negatively predict RAPM task performance provides further support for the idea that the aggregate measure of strategic behavior may reflect additional sources of variance such as using both strategies during the task. That is, participants may have used both response elimination and constructive matching strategies throughout the entire RAPM task and the questions that they were asked for each strategy type were asked at the very end with respect to the entire task. As a result, participants may report using both strategies in the post-experimental questionnaire. This positive bias may result from both strategies being used to complete different problems in the RAPM task, but a measure of intraindividual variability should be unrelated to this bias because it must necessarily be based on a more sensitive measure of differences in strategic behavior (the positive bias is due to the precision of the measure not an innate bias of the participant). The results of Experiment 1 did not support H1, indicating that individual differences in strategic behavior are not the reason that WMC and RAPM task performance share variance.

However, the data do support H2; the correlation between WMC and RAPM task performance was present in a control condition but absent in a constructive matching condition, indicating that individual differences in strategic behavior are an underlying reason that WMC and RAPM task performance correlate. This seeming contradiction is resolved by considering the validity of the strategy measures. The use of an aggregate measure of strategic behavior in the presence of theoretical predictions of intraindividual differences in strategic behavior across RAPM problems (e.g., Vigneau et al., 2006) renders an examination of the data for support in favor of H1 impossible in Experiment 1. As a result, Experiment 2 will examine strategic behavior across problems of increasing difficulty using more appropriate measures of strategic behavior generated to test H3 and H4, which parallel H1 and H2. Specifically, H3 states that intraindividual variability in strategic behavior will fully mediate the relation between WMC and RAPM task performance. Additionally, H4 states that the correlation between WMC and RAPM task performance will not be significant in a constructive matching condition (though WMC will be significantly related to RAPM task performance in a control condition). Thus, the goal of Experiment 2 was to replicate Gonthier and Thomassin (2015) using a measure of strategic behavior more sensitive to intraindividual variation in strategic behavior.

CHAPTER 7

EXPERIMENT 2 METHODS

Participants

A total of 290 participants were recruited from the introductory psychology research participation pool at Arizona State University. Nine participants had missing data on at least one of the tasks and were thus not included in the analyses. A total of 23 participants had less than 80% accuracy on the processing task and were excluded from analyses. The remaining sample was examined for any remaining outliers using mahalanobis distance outlier detection, which revealed one additional participant whose data were also excluded. This resulted in a final sample of 257 participants included in analyses. There were 125 participants in the constructive matching condition, and 132 participants in the control condition.

Procedure

All participants consented to participate in accordance with the standards of the Arizona State University's Institutional Review Board. After consenting to participate, all participants completed the following tasks in order: 1) Operation Span, 2) Reading Span, 3) Symmetry Span, 4) RAPM, and 5) a post-experimental questionnaire. The only task that was different than the version used in Experiment 1 was the RAPM task. All other tasks were exactly the same as previously described.

Materials

RAPM alterations. The RAPM task used in Experiment 1 was altered to include questions after each problem solved by the participant. Specifically, once a participant indicated their answer to a specific RAPM problem, they were asked the two constructive matching and two response elimination questions presented in *Table 1*. During the RAPM task, participants were not asked the filler questions. After the four strategy questions were randomly presented, participants were asked to describe their strategy in an untimed free-response question. Aside from the addition of these questions after each problem in RAPM, there were no changes to the task. The full version of the post-experimental questionnaire was still presented at the end of the experiment.

The measures of intraindividual variability in strategic behavior were calculated for each participant from their responses to the strategy questions they were asked to answer after each RAPM problem. This measure of strategic behavior represents variability in strategic behavior within a participant across RAPM problems. Responses to each of the strategy questions were summed across problems separately for each strategy question (e.g., problem 1, constructive matching question 1 strategy questionnaire response + problem 2, constructive matching question 1 strategy questionnaire response + . . . + problem 18, constructive matching question 1 strategy questionnaire response). This value for each strategy question was then divided by the total number of RAPM problems (18) to determine the mean reported use of a given strategy within a participant (as measured by a specific question) across all RAPM problems. The standard deviation for the participants' mean reported strategic behavior was then divided by the mean (the coefficient of variation) to calculate intraindividual variability in strategic behavior within a participant.

CHAPTER 8

EXPERIMENT 2 RESULTS

Experiment 2 was designed to conceptually replicate the findings in Gonthier and Thomassin (2015) showing that 1) individual differences in strategic behavior fully mediate the relation between WMC and RAPM accuracy (H3), and 2) when participants are provided with environmental support leading to the use of an effective strategy in a RAPM task, the correlation between WMC and RAPM accuracy is reduced (H4). To assess H4, all three proportional partial-unit span scores for the three WMC complex span tasks were submitted to a factor analysis using maximum likelihood estimation (and varimax rotation) and factor scores were derived. Descriptive statistics are presented for each of the variables in *Table* 4, and the correlations between the measures for both conditions appear in *Table* 5.

Table 4

Descriptive Statistics for the Conceptual Replication of Gonthier & Thomassin (2015)

	Control	Constructive Matching
Task	M (SD)	M (SD)
Operation Span Partial-Unit Span Score	0.80 (0.17)	0.79 (0.18)
Reading Span Partial-Unit Span Score	0.77 (0.17)	0.77 (0.19)
Symmetry Span Partial-Unit Span Score	0.64 (0.21)	0.66 (0.20)
Span Factor Score	-0.01 (0.82)	0.01 (0.89)
RAPM Accuracy	0.58 (0.18)	0.61 (0.19)
Response Elimination Q1	6.05 (2.52)	6.36 (2.43)
Response Elimination Q2	6.35 (2.35)	6.58 (2.23)
Constructive Matching Q1	7.59 (1.78)	8.04 (1.23)
Constructive Matching Q2	7.64 (1.73)	7.94 (1.58)
Strategy Coverage	7.03 (1.84)	7.28 (1.83)
Constructive Matching Q1 Coefficient of Variation	0.21 (0.18)	0.19 (0.18)
Constructive Matching Q2 Coefficient of Variation	0.26 (0.18)	0.22 (0.18)
Response Elimination Q1 Coefficient of Variation	0.36 (0.29)	0.32 (0.27)
Response Elimination Q2 Coefficient of Variation	0.36 (0.31)	0.30 (0.27)

Table 5

Correlations in tbe Constructive Matching (bottom triangle) and Control (top triangle) conditions

Variable No.	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Operation Span Partial-Unit Span Score	--	**0.57**	**0.24**	**0.84**	0.10	0.09	0.10	0.00	0.03	0.00	-0.08	-0.10	-0.15	-0.10
2	Reading Span Partial-Unit Span Score	**0.53**	--	**0.35**	**0.90**	**0.25**	-0.01	0.03	0.07	0.02	0.03	-0.11	-0.05	-0.05	-0.06
3	Symmetry Span Partial-Unit Span Score	**0.38**	**0.32**	--	**0.49**	**0.18**	-0.12	-0.10	0.04	0.07	0.11	0.02	0.02	0.08	0.07
4	Span Factor Score	**0.83**	**0.89**	**0.52**	--	**0.22**	0.01	0.04	0.04	0.04	0.03	-0.10	-0.07	-0.08	-0.07
5	RAPM Accuracy	0.13	**0.25**	**0.24**	**0.25**	--	-0.07	-0.06	**0.18**	0.11	0.15	-0.06	-0.01	**0.21**	**0.22**
6	Response Elimination Q1	-0.07	-0.09	-0.15	-0.11	-0.09	--	**0.57**	**0.18**	0.17	0.13	**-0.26**	**-0.45**	**-0.66**	**-0.68**
7	Response Elimination Q2	-0.09	-0.03	-0.11	-0.08	-0.11	**0.69**	--	**0.21**	**0.18**	-0.02	**-0.18**	**-0.34**	**-0.54**	**-0.60**
8	Constructive Matching Q1	**0.23**	**0.35**	**0.20**	**0.35**	**0.39**	0.11	0.13	--	**0.51**	**0.32**	**-0.31**	**-0.21**	0.06	0.01
9	Constructive Matching Q2	-0.02	0.10	-0.04	0.05	0.16	**0.21**	**0.18**	**0.38**	--	**0.27**	**-0.18**	**-0.19**	0.11	0.13
10	Strategy Coverage	0.07	0.09	0.12	0.11	**0.25**	0.11	0.09	**0.37**	**0.32**	--	**-0.23**	-0.15	0.02	-0.02
11	Constructive Matching Q1 Coefficient of Variation	0.07	-0.11	0.08	-0.02	-0.06	**-0.27**	**-0.29**	**-0.30**	**-0.36**	**-0.22**	--	**0.82**	**0.45**	**0.49**
12	Constructive Matching Q2 Coefficient of Variation	0.07	-0.05	0.05	0.01	-0.01	**-0.33**	**-0.34**	**-0.25**	**-0.42**	**-0.28**	**0.90**	--	**0.66**	**0.65**
13	Response Elimination Q1 Coefficient of Variation	0.15	0.06	0.13	0.12	**0.18**	**-0.57**	**-0.53**	-0.02	-0.16	-0.07	**0.64**	**0.69**	--	**0.92**
14	Response Elimination Q2 Coefficient of Variation	0.07	0.02	0.15	0.07	0.13	**-0.52**	**-0.54**	-0.06	**-0.21**	-0.09	**0.59**	**0.66**	**0.91**	--

Note: Significant correlations are bolded. Colors are discussed in detail in the text.

The variables in *Table* 4 were submitted to independent-samples t-tests with the strategy condition manipulation as a between-subjects variable. There was a significant effect of strategy condition on responses to the first constructive matching question, $t(255) = 2.341$, $p = .020$. Participants reported using constructive matching more in the constructive matching condition than the control condition. This indicates that the constructive matching manipulation was effective in increasing use of a constructive matching strategy. However, there were no differences in the reported use of response elimination across strategy conditions. The strategy manipulation was designed to increase the use of a constructive matching strategy. Although it is interesting that increasing reported use of constructive matching is not also associated with a decrease in the use of response elimination, not finding differences in response elimination across conditions does not provide any information about the effectiveness of a constructive matching manipulation. Finding differences in reported constructive matching strategic behavior is sufficient and indicates that the manipulation was effective. There were no other significant differences across strategy conditions in any of the other measures in *Table* 4. The absence of an effect of the strategy manipulation on RAPM task performance may be related to the fact that only low working memory capacity participants appeared to benefit from the strategy manipulation in Experiment 1.

Similar to H2 in Experiment 1, H4 predicts that the *relation* between WMC and RAPM task performance will be reduced in the constructive matching condition. Thus, to evaluate this hypothesis the correlations in *Table* 5 were examined further. There are several interesting things to note about the correlations in *Table* 5. First, the working memory span scores all correlate with each other and the factor score representing the shared variance across all three of the complex span tasks (correlations in yellow). Second, many of the retrospective strategic behavior reports were still positively correlated as in Experiment 1 indicating that there may still be a bias to respond positively across all aggregate retrospective strategy questions (correlations in orange). The intraindividual variability in strategic behavior variables were also positively correlated in both conditions (correlations in pink). This indicates that participants who were more variable in the use of a constructive matching strategy tended to be more variable in the use of a response elimination strategy. The intraindividual variability variables were also negatively related to most of the retrospective strategy reports (correlations in purple). This indicates that more variable strategy reports by participants tended to relate to less of either type of strategy being reported.

Third, the WMC factor score was significantly correlated with RAPM performance in the control condition (*N* = 132, *r* = .22, *p* = .01) and in the constructive matching condition (*N* = 125, *r* = .247, *p* = .006). Importantly, unlike the results of Experiment 1 the correlation between RAPM task performance and WMC was not significantly different across strategy conditions, Fisher r-to-z = 0.19, *p* = .85 (correlations in green). Thus, these data do not support H4 in Experiment 2 (see *Figure* 10). This raises concern that asking participants to report their strategic behavior after each question may have also changed how participants approached the entire experiment.[5] As shown in *Figure* 10, low working memory capacity participants did not

---

[5] The data were split into quartiles based on WMC and the upper quartile formed a high WMC group while the lower quartile formed a low WMC group. RAPM accuracy was submitted to a three-factor ANOVA with Experiment (1 v.s. 2), Strategy Condition (Constructive Matching v.s. Control), and Span (Low v.s. High) as between-subjects factors. The interaction between Experiment, Strategy Condition, and Span was significant, *F*(1, 271) = 4.749, *p* = .030, *MSE*= 0.029, partial $\eta 2$ = .017.

experience a benefit from the constructive matching condition as they did in Experiment 1[6]. This might be expected if asking participants to report strategic behavior after each RAPM problem increased the use of constructive matching strategic behavior (similar to Experiment 4 in Loesche et al., 2015).

---

[6] RAPM accuracy for low WMC participants in Experiment 1 was significantly better for participants in a constructive matching ($M = .54$, $SD = .13$) compared to a control ($M = .42$, $SD = .17$) condition, $F(1, 73) = 10.303$, $p = .002$, $MSE = 0.022$, partial $\eta 2 = .124$. However, RAPM accuracy for low WMC participants in Experiment 2 was similar for participants in a constructive matching ($M = .54$, $SD = .18$) compared to a control ($M = .51$, $SD = .18$) condition, $F(1, 62) = 0.357$, $p = .552$, $MSE = 0.034$, partial $\eta 2 = .006$.

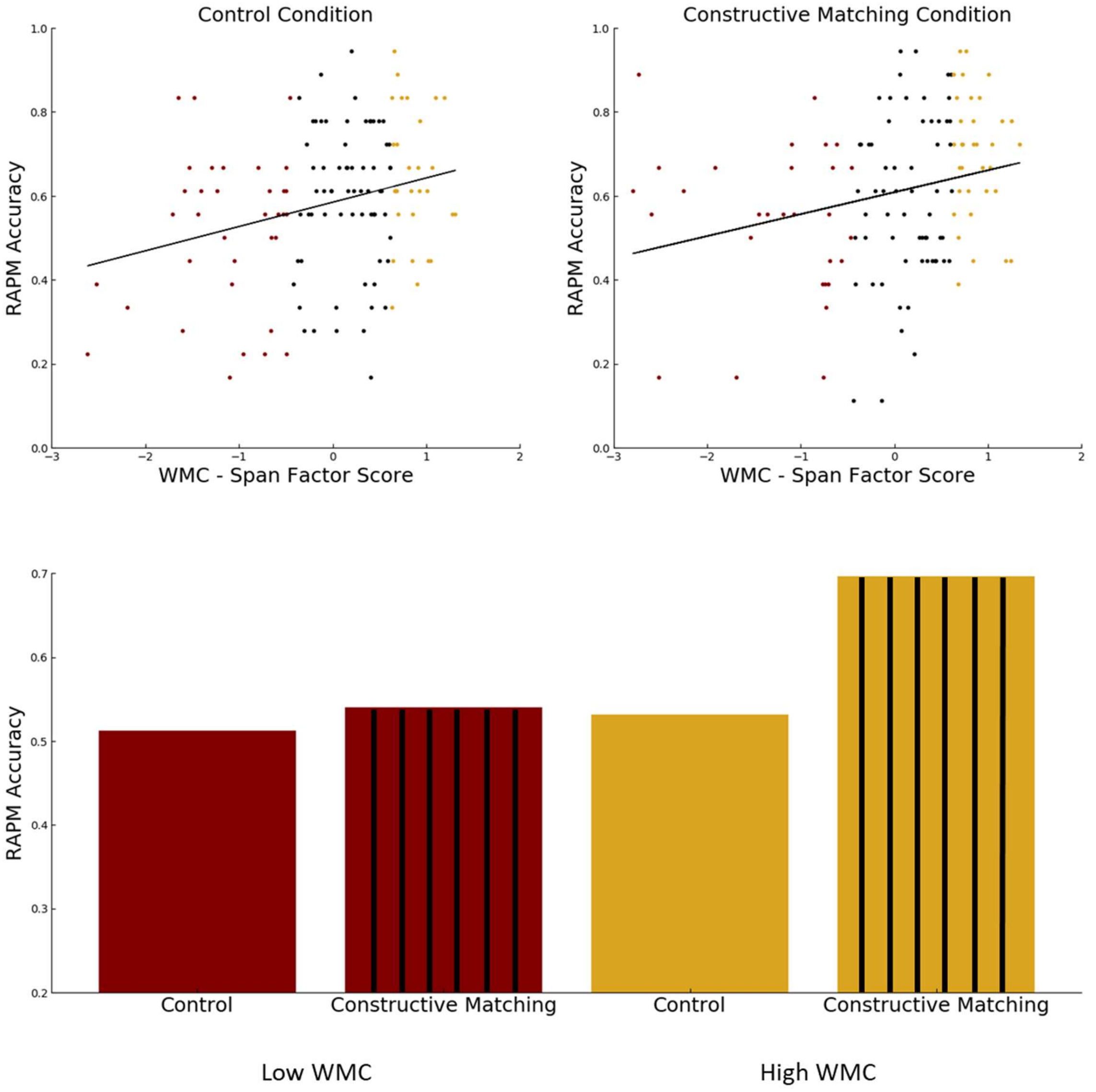


Figure 10. Top: Relation between WMC and RAPM accuracy for participants in the Control (left) and Constructive Matching (right) conditions. Bottom: Mean RAPM accuracy for low WMC and high WMC participants in the Control and Constructive Matching conditions. Described in detail in the text.

In line with the predictions following from Experiment 4 in Loesche et al. (2015), examining RAPM accuracy of low working memory capacity participants across experiments in *Figures* 8 and 10 illustrates that asking participants to report their strategies after each question facilitated the performance of low working memory capacity participants in the control condition[7] and thus likely reduced the potential to observe an improvement in RAPM performance due to the constructive matching manipulation. Interestingly, the performance of high working memory capacity participants was also affected by the embedded reports of strategic behavior throughout the task (though not enough to alter the relation between WMC and RAPM task performance across conditions in Experiment 2). Consistent with Loesche et al. (2015), these participants appear to benefit from the constructive matching manipulation[8].

However, it is worth noting that comparing *Figures* 8 and 10 reveals a potential decrease in RAPM accuracy for high working memory capacity participants in Experiment 2 in the control condition[9] and a potential increase in accuracy in Experiment 2 in the constructive matching condition[10] (compared to the data presented in *Figure* 8 in Experiment 1). Thus, H4 cannot be evaluated in Experiment 2 because the method of assessing strategic behavior appears to have changed how participants complete the RAPM task. This change is potentially driven by the same increase in constructive matching strategic behavior observed in Loesche et al. (2015), though the effect on constructive matching strategic behavior appears to vary as a function of both span and

---

[7] The effect of Experiment (1 vs. 2) on RAPM accuracy was marginally significant for low WMC participants in the Control condition, $F(1, 58) = 3.776$, $p = .057$, $MSE = 0.032$, partial $\eta 2 = .061$. Accuracy was higher in Experiment 2 ($M = .51$, $SD = .18$) compared to Experiment 1 ($M = .42$, $SD = .17$), indicating that asking low WMC participants to report their strategic behavior after each problem improved their performance on the RAPM task.

[8] The effect of Strategy Condition on RAPM accuracy for high WMC participants in Experiment 2 failed to reach significance, $F(1, 62) = 2.997$, $p = .088$, $MSE = 0.028$, partial $\eta 2 = .046$. However, there was a trend for participants in the constructive matching condition to have higher accuracy ($M = .70$, $SD = .15$) than participants in the control condition ($M = .62$, $SD = .17$).

[9] There were no differences in RAPM accuracy across Experiments 1 and 2 for high WMC participants in the control condition, $p = .835$.

[10] RAPM accuracy was higher in Experiment 2 ($M = .70$, $SD = .15$) compared to Experiment 1 ($M = .59$, $SD = .17$) for high WMC participants in the constructive matching condition.

strategy condition in the present study. The possible increase in the use of constructive matching strategic behavior for low span participants in the control condition (e.g., Loesche et al., 2015) makes it difficult to evaluate H4 in Experiment 2.

The average retrospective report of constructive matching strategic behavior was submitted to an exploratory two-factor ANOVA with Experiment (1 vs. 2) and Condition (Constructive Matching vs. Control) as between-subjects factors and WMC entered as a covariate. There was a main effect of WMC, $F(1, 552) = 7.215$, $p = .007$, $MSE= 1.838$, partial $\eta 2 = .013$, indicating that high WMC participants were more likely to use a constructive matching strategy. Additionally, the average retrospective report of constructive matching strategic behavior was lower in the control condition ($M = 7.474$, $SD = 2.10$) compared to the constructive matching condition ($M = 7.817$, $SD = 1.82$) across both experiments combined, $F(1, 552) = 8.405$, $p = .004$, $MSE= 1.838$, partial $\eta 2 = .015$. This indicates that after statistically controlling for WMC, the experimental manipulation was effective across both experiments in increasing the reported use of constructive matching strategic behavior.

Additionally, participants reported using constructive matching more in Experiment 2 ($M = 7.805$, $SD = 2.01$) compared to Experiment 1 ($M = 7.486$, $SD = 1.96$) across both strategy conditions combined, $F(1, 552) = 7.251$, $p = .007$, $MSE= 1.838$, partial $\eta 2 = .013$. These findings are similar to the findings in Experiment 4 in Loesche et al. (2015) demonstrating that providing the appropriate rules to participants increased the use of a constructive matching strategy. Although the two-factor ANOVA conducted in the present study was exploratory, it is possible that there is a similar underlying process explaining findings in the present study and findings in Loesche et al. (2015).

Specifically, Loesche et al. (2015) provided rules to participants in an effort to understand how goal management processes influence the shared relation between WMC and gF when the appropriate rules were provided to participants (minimizing the need for rule induction). Providing the rules was associated with an increase in the relation between WMC and gF, indicating that the rule induction process (or whatever process was influenced by providing the rules to participants) was a nuisance variable and not an important underlying process responsible for the shared variance between WMC and gF. Although the correlation between WMC and gF was the same across conditions in Experiment 2, there was no interaction between Experiment and Condition, $p = .815$, indicating that the effect of the strategy manipulation on reported constructive matching strategic behavior was similar across Experiments 1 and 2.

However, the reduced correlation between WMC and gF in the constructive matching condition in Experiment 1 was not observed in Experiment 2, which may have been due to an increase in reported constructive matching behavior in Experiment 2 across both strategy conditions. That is, differences in the data presented in *Figure* 10 compared to *Figure* 8 appear to result from an increase in RAPM task performance (likely driven by an increase in the use of a constructive matching strategy) for low working memory capacity participants in the control condition and an increase in RAPM task performance (likely driven by an increase in the use of a constructive matching strategy) for high working memory capacity participants in the constructive matching condition. Collectively, this may indicate that inferring which strategy is most effective for a given RAPM problem requires goal maintenance. By minimizing the need to weigh support for the effectiveness of each strategy by allowing time to reflect on strategic behavior, low WMC participants were able to achieve performance equivalent to their performance in a constructive matching condition.

This indicates that low WMC participants may occasionally fail to select strategic behavior consistent with their goals in a traditional RAPM task possibly due to failures in goal maintenance. Participants should consider all rules collectively to arrive at the correct solution in an RAPM task and needing to consider the appropriate strategic behavior may have been less salient of a goal for low WMC participants if the information held in primary memory was already the amount of information that the participant was capable of holding in primary memory (primary memory capacity). In Experiment 2 low WMC participants had RAPM task performance comparable to their performance in the constructive matching condition, and exploratory analyses indicated that participants were more likely to use constructive matching in Experiment 2 compared to Experiment 1.

The goal of considering the appropriate strategic behavior was attained in Experiment 2 by providing additional time and asking questions that required participants to reflect on their strategic behavior after each problem on the RAPM task. This may have allowed more information pertinent to solving the problem to be stored in primary memory which should result in better performance for problems that require maintenance of more information than could have previously been held in primary memory when the goal to select the appropriate strategy needed to be held in primary memory as well. Problems that require maintenance of more information than can be held in primary memory even after the appropriate strategic behavior has been selected should be associated with lower accuracy for that problem. The low WMC participants in

Experiment 2 did not need to hold the goal to consider the most appropriate strategy in primary memory during each RAPM problem.

Although this may have resulted in higher performance in the control condition due to the maintenance of more information pertinent to the problem in primary memory, by encouraging selection of the constructive matching strategy the goal of selecting the appropriate strategy was supported by the task environment. More information pertinent to the problem may be maintained in primary memory for low WMC participants in the constructive matching condition for the same reason that more information could be maintained in primary memory for low WMC participants in the control condition. The need to select strategic behavior consistent with goals during an RAPM task was achieved by asking participants to reflect on strategic behavior after each problem and may have allowed for more information relevant to the problems to be stored in primary memory due to the reduced need to actively maintain that goal in memory. However, any RAPM problems requiring consideration of more information than can be held in primary memory will still be associated with reduced accuracy on that problem which may explain why the constructive matching manipulation in the present study was unable to improve RAPM task performance above the improvement seen due to a reduced need to maintain information about the goal to select the most appropriate behavior for that problem. This may reflect capacity limitations of low WMC participants.

High WMC participants did not benefit from frequent reminders to consider strategic behavior in the control condition. This indicates that the goal to select the appropriate strategy may have already been held in primary memory for high WMC participants in Experiment 1. Critically, this would imply that part of the reason that high and low WMC participants exhibit differences in RAPM task performance is due to a failure to select strategic behavior consistent with task goals. Further, these failures in goal maintenance should arise when the amount of information to be held in primary memory is at or exceeds capacity and the goal to reflect on the appropriate strategy may not be stored in primary memory, which may lead to more response elimination strategic behavior as exhibited by low WMC participants in Jarosz and Wiley (2012).

High WMC participants in the constructive matching condition in Experiment 2 performed better on the RAPM task when both the structure of the RAPM problem supported the use of constructive matching and strategic behavior was reflected on after each RAPM problem, indicating that either or both the manipulation in Experiment 1 and the time to consider appropriate strategic behavior in Experiment 2 reduced the amount of information required for goal maintenance that needs to be represented in primary memory. Thus, part of the reason that WMC shares variance with RAPM task performance may be differences in strategic behavior resulting from capacity limitations and differences in goal maintenance (attention control). If selecting the response elimination strategy is less efficient than selecting a constructive matching strategy, then more information would need to be maintained when response elimination strategic behavior is exhibited which could also contribute to lower RAPM accuracy.

The loss of more information from primary memory due to an increase in the amount of information that must be simultaneously held in memory to perform response elimination compared to constructive matching should result in more reliance on cue-dependent retrieval of information from secondary memory when a response elimination strategy is used rather than a constructive matching strategy. Collectively, these predictions are consistent with Gonthier and Thomassin (2015) and Unsworth et al. (2014) and the similar correlations across strategy conditions in Experiment 2 further indicate that providing environmental support for the use of a constructive matching strategy and reflecting on strategic behavior after each RAPM problem provided only a partial explanation for the shared variance between WMC and RAPM task performance.

If the reports of strategic behavior after each problem were associated with an increase in the relation between WMC and gF compared to a single aggregate measure of strategic behavior, then goals relevant to selecting the appropriate strategy by considering strategic behavior after each RAPM problem interfere with the ability to observe shared variance and represent a nuisance variable. The absence of repetitive strategic thinking may tax goal maintenance processes and reduce how much of the problem can be maintained using a constructive matching strategy. When repetitive inferences are made about strategic behavior, goal maintenance processes specific to the problem being solved may more adequately reflect processes important for solving a RAPM problem. If the constructive matching strategy condition was associated with a decrease in the relation between WMC and gF compared to the control condition, then manipulating the amount of support for the use of a constructive matching strategy indicates that the strategic behavior implemented is an important mechanism underlying the relation between WMC and RAPM task performance. Due to the opposing expected effects on the relation between WMC and RAPM task performance, there may be no overall change in the relation between WMC and RAPM task performance in the constructive matching condition compared to the control condition in Experiment 2.

The goal maintenance processes specific to evaluating strategic behavior for each RAPM problem along with the reduction in environmental support for the use of response elimination may have resulted in the positive correlation between WMC and gF in the constructive matching condition in the present study which did not differ from the correlation in the control condition. The goal maintenance processes specific to evaluating strategic behavior for each problem may be a nuisance variable due to capacity limitations, and may result in selection of a response elimination strategy for at least part of the problem out of necessity because constructive matching strategic behavior is not possible if the amount of information held in primary memory exceeds the information needed to construct the solution to the problem though the participant is aware of the more appropriate strategy. Vigneau et al. (2006) indicated that not all of the variance in strategic behavior in an RAPM task overlaps with WMC. Thus, it is possible that primary memory capacity will predict response elimination strategic behavior and predict performance on the RAPM task.

The two response elimination coefficients of variation were correlated with RAPM task performance in the control condition, and the response elimination coefficient of variation for the first response elimination question was correlated with RAPM task performance in the constructive matching condition. However, more variability in response elimination strategic behavior was associated with better RAPM task performance, again raising concern that asking strategy questions after each problem changed the nature of how participants completed the task. The two constructive matching coefficients of variation were unrelated to RAPM task performance in either condition (correlations in gray). Further examination of the correlations in *Table* 5 again highlights the likelihood that the data will not support H3 (or H1). The WMC factor score was not correlated with any of the old or new measures of strategic behavior with one exception (correlations in blue). WMC was positively related to constructive matching retrospective strategy reports for the first constructive matching strategy question in the constructive matching condition only, $p < .01$. Consistent with previous exploratory analyses, high working memory capacity participants were more likely to report using a constructive matching strategy in the constructive matching condition.

However, to assess H3 directly a structural model was fit to the data from the control condition, as in Experiment 1. Given the concern that the strategy reports in Experiment 2 changed the way the task was completed, first models using the same variables for strategic behavior as Experiment 1 were fit to the data followed by models using intraindividual variability in strategic behavior as mediators. H3 states that intraindividual variability in strategic behavior will fully mediate the relation between WMC and gF. Experiment 1 measures of strategic behavior were likely noisy due to intraindividual variation in strategic behavior paired with a single strategic behavior report. Although the measures of strategic behavior in Experiment 2 (coefficient of variation) were designed to address this limitation, the way that participants complete a RAPM task is by asking participants to reflect on strategic behavior after each RAPM problem. Thus, these results should be interpreted with caution with respect to H3.

A full structural model was fit to the data from the control condition in MPLUS (Muthén & Muthén, 1998-2011) using the same retrospective strategic behavior variables as Experiment 1 (see *Figure* 11). The bias to respond positively to using both strategies due to the retrospective nature of the questions was controlled for by allowing the strategy variables to correlate outside of the model. The model in *Figure* 11 provided a good fit to the data, χ^2 (16) = 16.606, p = .41. An examination of the fit indices in indicated that the model was able to recreate the sample correlation matrix over 94% of the time (RMSEA = .017 and SRMR = .056). Additionally, both the CFI (0.996) and TLI (0.994) indicated that the model provided an improvement over a baseline model that fixes the correlations across latent factors and measured variables to 0.

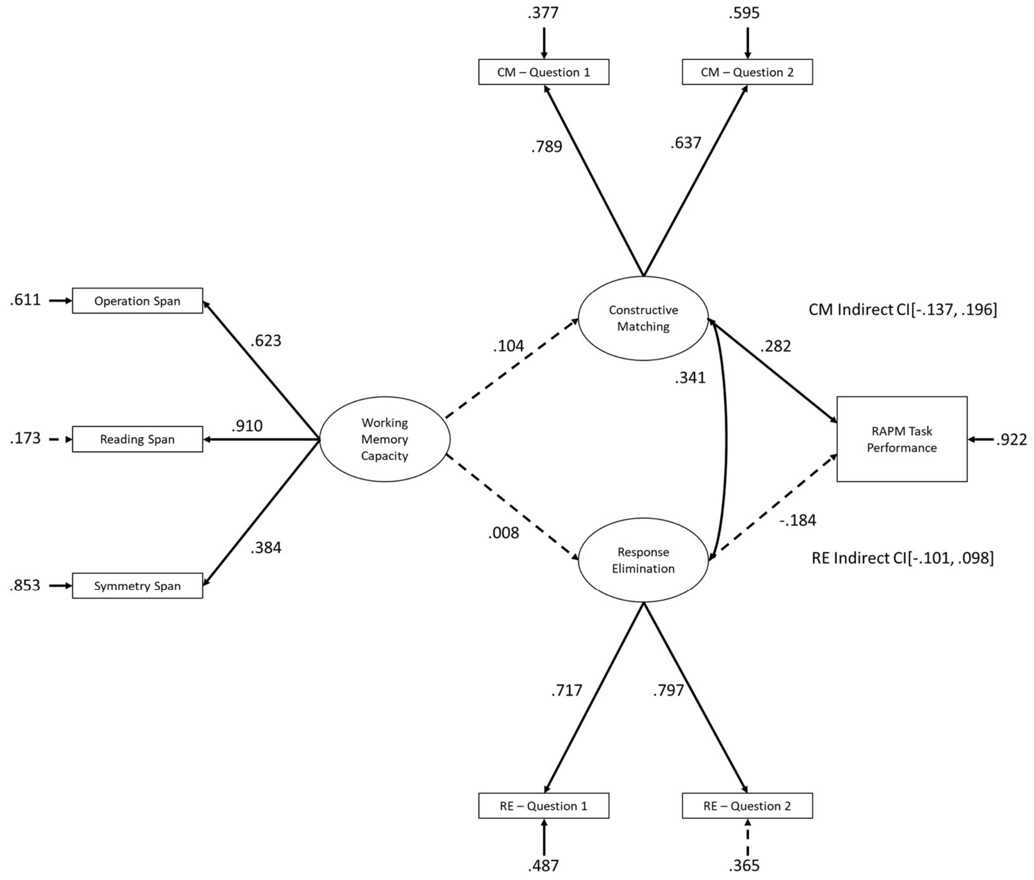


Figure 11. Structural model for Experiment 2 control condition data. Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

Despite the close fit of the model to the data, the relation between WMC and RAPM task performance via response elimination (standardized 95% bootstrapped CI[-0.101, 0.098]), and via constructive matching (standardized 95% bootstrapped CI[-0.137, 0.196]) was not significant. These indirect effects may not be present partially because the nature of the task may have changed. As shown in *Figure* 11, WMC did not predict retrospective reports of constructive matching or response elimination strategic behavior. Retrospective reports of constructive matching strategic behavior were positively related to RAPM task performance. However, as indicated in *Figure* 11 (and consistent with Experiment 1 results), individual differences in the reported use of a response elimination strategy did not predict RAPM task performance, though the regression coefficient is negative as would be predicted (Bethel-Fox et al., 1984; Vigneau et al., 2006; Gonthier & Thomassin, 2015). Although this is inconsistent with H3, these measures were only presented for the sake of comparison due to the apparent change in the way participants complete the RAPM task (see comparisons of *Figures* 8 and 10). A structural model was estimated using the coefficients of variation in strategic behavior as mediators instead of retrospective reports of strategic behavior. A latent factor representing variability in constructive matching strategic behavior was estimated from the coefficients of variation for the two constructive matching questions. Additionally, a latent factor representing variability in response elimination strategic behavior was estimated from the coefficients of variation for the two response elimination questions.

The full structural model was fit to the data. However, there appeared to be linear dependencies in the coefficients of variation for the constructive matching questions. As a result, the two constructive matching and two response elimination coefficient of variation variables were averaged, and these averages were used rather than latent factors in the mediation model presented in *Figure* 12. The model in *Figure* 12 provided a reasonable fit to the data, χ^2 (7) = 13.781, *p* = .055. An examination of the fit indices illustrates that the model was able to recreate the sample correlation matrix over 91% of the time (RMSEA = 0.086 and SRMR = 0.070). Additionally, both the CFI (0.952) and TLI (0.897) indicated that the model provided > 88% improvement over a baseline model that fixes the correlations across all variables to 0. Despite the close fit of the model to the data, the relation between WMC and RAPM task performance via intraindividual variability in constructive matching (standardized 95% bootstrapped CI[-0.031, 0.082]), and via intraindividual variability in response elimination (standardized 95% bootstrapped CI[-0.103, 0.047]) was not significant. This could be due to the additional requests to consider strategic behavior throughout the task. Thus, although H3 and H4 were not supported in Experiment 2, it is unclear whether the design of the experiment changed the nature of the RAPM task or the relation between WMC and RAPM task performance.

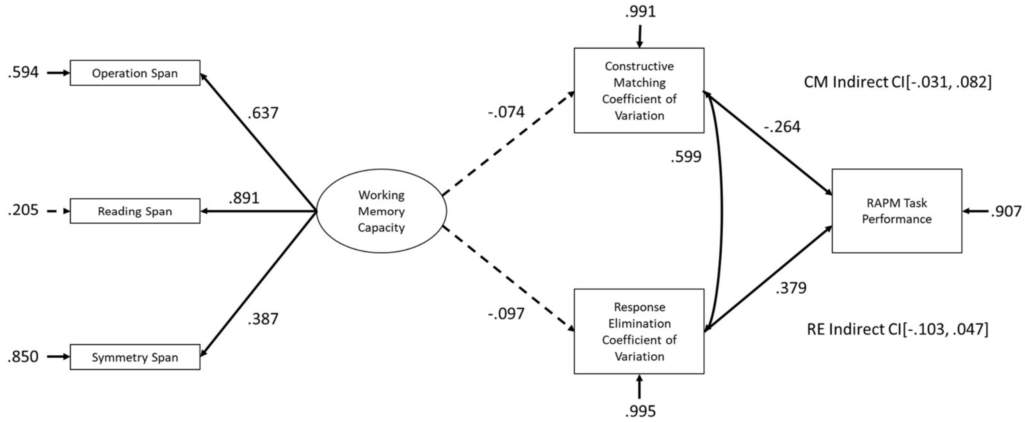


Figure 12. Structural model for Experiment 2 control condition data. Latent factor variance is fixed to 1 and latent factor mean is fixed to 0. Described in detail in the text.

More variability in the application of constructive matching and response elimination strategies should be associated with worse RAPM task performance and lower WMC, whereas less variability in the application of these strategies should be associated with better RAPM task performance and higher WMC. That is, the higher ability participants should be able to successfully implement a constructive matching strategy consistently, whereas a low ability participant may be more variable in how well they are able to implement constructive matching strategic behavior across problems of increasing difficulty. The coefficients of variation in strategic behavior did not relate to WMC in the model presented in *Figure* 12, and less variability in constructive matching strategic behavior was indeed related to better RAPM task performance. However, more variability in response elimination strategic behavior was associated with better RAPM task performance. This may indicate that being able to flexibly use response elimination across problems is associated with better RAPM task performance.

Although there was no support for H3 or H4 in Experiment 2, the measure of strategic behavior in the RAPM task in Experiment 2 appears to have changed the nature of the relationship between WMC and gF. Low WMC participants were more accurate when they were prompted to reflect on their strategic behavior throughout the task, and high WMC participants were more accurate in a constructive matching condition when they were prompted to reflect on their strategic behavior throughout the task. As a result, rather than attempting to increase the sensitivity of the measure of strategic behavior, power concerns were addressed in Experiment 3 with a larger sample size. In Experiment

3, a subset of data was analyzed that were collected for a large scale individual differences study conducted as a part of an NSF grant[11].

Multiple measures of each cognitive and strategic process were collected, as well as multiple measures of WMC and gF. Participants were asked to answer the four strategy questions in Experiment 1 and Experiment 2 after completing the gF measures in Experiment 3 (the four questions were asked retrospectively as they pertained to each of the gF tasks). Although these strategy measures were aggregate measures (the gF tasks could not be edited to maintain the integrity of the data collected for the grant), the sample size for Experiment 3 should offset the power concerns that arise from the use of these types of aggregate measures. The goal of Experiment 3 is to determine the mechanisms underlying the relation between WMC and gF. To achieve this, two hypotheses were evaluated based on data collected for NSF grant 1632327. According to H5, cognitive processes will fully mediate the relation between WMC and gF. By contrast, H6 states that strategic behavior will no longer mediate the relation between WMC and gF after controlling for individual differences in cognitive processes. The model in *Figure* 5 was fit to the data in Experiment 3 to assess these hypotheses and the predictions outlined in the introduction.

---

[11] Data were still being collected at the time this manuscript was written.

CHAPTER 9

EXPERIMENT 3 METHODS

Participants

Experiment 3 was based on a subset of the data collected for NSF grant 1632327. A total of 974 participants completed the study. Data were collected at three different Universities: 1) University of Oregon *N* = 187, 2) Purdue University *N* = 406, 3) Arizona State University *N* = 381. A total of 51 participants had missing data on at least one of the tasks and were excluded from analyses. Eight participants that made errors on over half the trials in the working memory complex span tasks were excluded from analyses[12]. An additional 15 participants did not complete the arrays tasks correctly, and 20 participants did not complete the source memory for pictures task correctly. These participants were excluded from analyses. There were an additional five participants that were excluded from analyses who had zero correct on at least two lists in the immediate free recall task, 33 participants were excluded due to lack of engagement during the cued paired associate task, and 34 participants were excluded for lack of engagement during the Flanker task. One participant exhibited excessive early presses in the psychomotor vigilance task (> 25), and two participants skipped through the questions in the intelligence tasks (< 30 seconds to complete the task). These participants were also excluded from analyses. Thus, data for 805 participants were included in analyses (University of Oregon *N* = 130, Purdue University *N* = 360, Arizona State University *N* = 315).

---

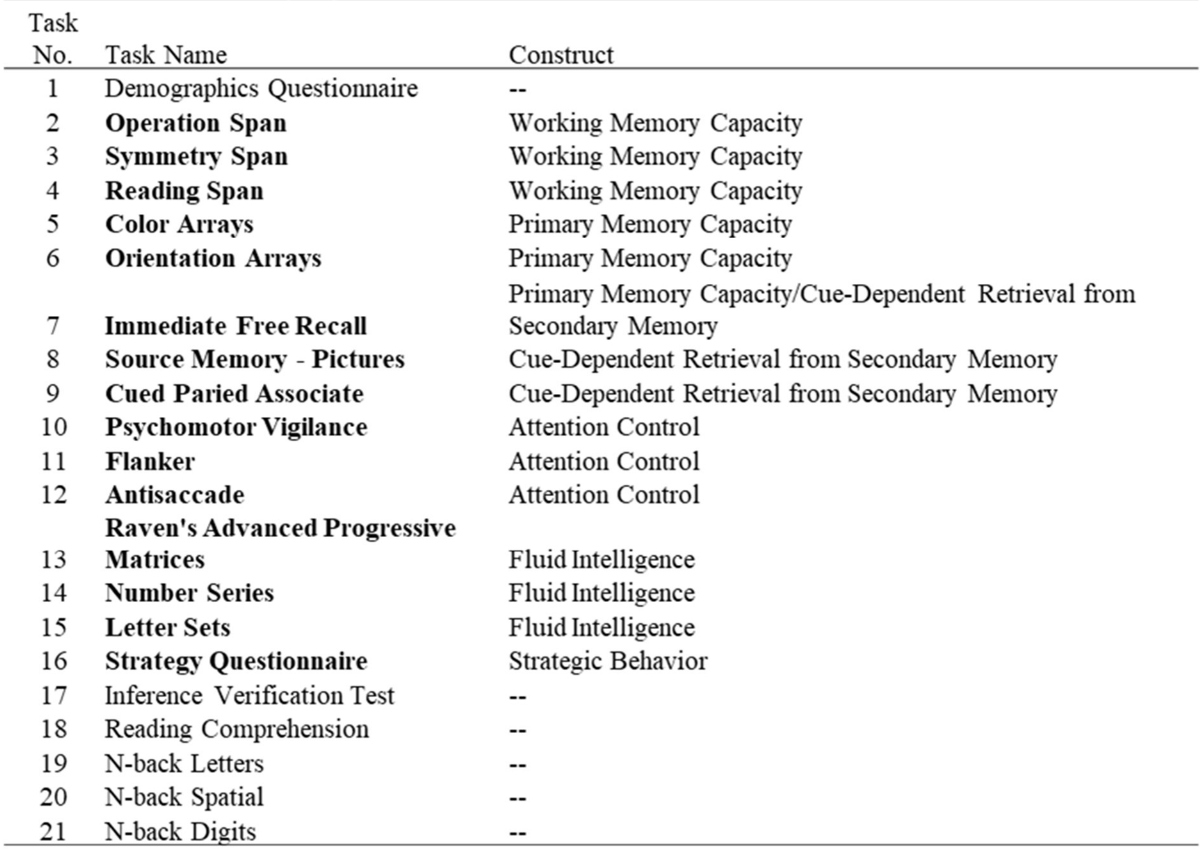
[12] The exclusion criteria for Experiment 3 were identical to the exclusion criteria used elsewhere for the NSF grant data.

Procedure

Relevant data were selected from a large individual differences study being conducted as a part of a grant addressing a completely different research question. The investigators at each institution agreed to include the strategy questions after the gF tasks were completed by participants. All participants consented to participate in accordance with the standards of the Institutional Review Board at each university. *Table* 7 contains the full list of tasks given to participants, who completed each task in the order indicated in the table. The tasks that are in bold in the table are the tasks participants completed that are relevant for the present study, and only data from these tasks were examined in Experiment 3. The construct each task represents is presented in the table as well for the tasks in bold.

Table 7

Tasks Included in the Study Experiment 3 Data Were Selected From



Task No.	Task Name	Construct
1	Demographics Questionnaire	--
2	**Operation Span**	Working Memory Capacity
3	**Symmetry Span**	Working Memory Capacity
4	**Reading Span**	Working Memory Capacity
5	**Color Arrays**	Primary Memory Capacity
6	**Orientation Arrays**	Primary Memory Capacity
7	**Immediate Free Recall**	Primary Memory Capacity/Cue-Dependent Retrieval from Secondary Memory
8	**Source Memory - Pictures**	Cue-Dependent Retrieval from Secondary Memory
9	**Cued Paried Associate**	Cue-Dependent Retrieval from Secondary Memory
10	**Psychomotor Vigilance**	Attention Control
11	**Flanker**	Attention Control
12	**Antisaccade**	Attention Control
13	**Raven's Advanced Progressive Matrices**	Fluid Intelligence
14	**Number Series**	Fluid Intelligence
15	**Letter Sets**	Fluid Intelligence
16	**Strategy Questionnaire**	Strategic Behavior
17	Inference Verification Test	--
18	Reading Comprehension	--
19	N-back Letters	--
20	N-back Spatial	--
21	N-back Digits	--

Materials

Working memory capacity. As in Experiments 1 and 2, working memory capacity was estimated from the partial-unit span scores from operation, reading, and symmetry span. List lengths for operation and reading span varied from 3 to 7, and list lengths in symmetry span varied from 2 to 5. Each list length was presented twice for 50 total possible points for operation and reading span and 28 total possible points for symmetry span. The estimate of working memory capacity is represented by the working memory capacity latent variable in *Figure* 5.

Primary memory capacity. Primary memory capacity was estimated from dependent variables from the color arrays, orientation arrays, and immediate free recall tasks. The individual tasks are described below and the estimate of primary memory capacity is represented by the primary memory capacity latent variable in *Figure* 5.

Color arrays. In the color arrays task (Morey & Cowan, 2004), four, six, or eight colored squares were presented briefly to participants. After a delay, the squares reappeared in the same location in either the same or a different color and one of the colored squares in this new display was circled. The participant was asked to indicate whether the color of the circled square matched the color of the square previously presented in that location. The dependent variable in this task was the bias-corrected measure of capacity, *k* (Cowan et al., 2005).

Orientation arrays. In the orientation arrays task (Luck & Vogel. 1997), five or seven colored rectangles were presented briefly to participants. After a delay, the rectangles appeared in the same location and color in either the same or a different orientation and one of the colored rectangles in this new display was circled. The participant was asked to indicate whether the orientation of the circled rectangle matched the orientation of the rectangle previously presented in that location and color. The dependent variable in this task was the bias-corrected measure of capacity, *k* (Cowan et al., 2005).

Immediate free recall – primary memory. In the immediate free recall task (Unsworth, Spillers, & Brewer, 2010), participants were presented with a list of ten words and asked to recall the words from the current list in any order upon seeing a '???' on the screen. A total of ten lists were presented to participants and the dependent variable was a primary memory estimate following Tulving and Colotla (1970).

Cue-dependent retrieval from secondary memory. Cue-dependent retrieval from secondary memory was estimated from dependent variables from the immediate free recall, source memory for pictures, and cued paired associate tasks. The individual tasks are described below and the estimate of cue-dependent retrieval from secondary memory is represented by the cue-dependent retrieval from secondary memory latent variable in *Figure* 5.

Immediate free recall – secondary memory. A secondary memory estimate was calculated from the immediate free recall task following Tulving and Colotla (1970).

Source memory – pictures. In the source memory task, participants were presented with pictures that appeared in one of four quadrants on the screen (Unsworth & Brewer, 2009). Participants were then presented with a set of pictures and asked to indicate whether the picture was old (previously seen) or new (was not previously studied). If the participant indicated the picture was old, they were further asked to indicate the quadrant that they believed the picture appeared in. The dependent measure in this task was the mean accuracy for the picture source (quadrant) judgements.

Cued paired associates. In the cued paired associates task (Brewer & Unsworth, 2012; Carpenter, Pashler, & Vul, 2006), participants studied lists of word pairs. Each word pair consisted of a cue and a target. After a list of word pairs were studied, participants were presented with cue words and asked to recall the target word from that pair. The dependent variable for this task was the number of target words correctly recalled in response to the cue words.

Attention control. Attention control was estimated from dependent variables from the psychomotor vigilance, flanker, and antisaccade tasks. The individual tasks are described below and the estimate of attention control is represented by the maintenance of goal-relevant information latent variable in *Figure* 5.

Psychomotor vigilance. In the Psychomotor Vigilance Task (PVT; Dinges & Powell, 1985), participants were presented with a timer set to zero (0.000) in the center of the screen. After a variable amount of time, the timer began counting up and the participant was asked to press the space bar as quickly as possible to stop the timer. The dependent variable was the mean response time from the slowest 20 percent of trials (Unsworth, Redick, Lakey, & Young, 2010).

Flanker. In the flanker task (e.g., Brewer & Unsworth, 2012), participants were asked to identify the direction of the center arrow (→→→→→) in a display of arrows. There were three types of stimuli in this task. On neutral trials, participants were asked to identify the direction of the center arrow which was surrounded by horizontal lines (- - → - -). On congruent trials participants were asked to identify the direction of the center arrow which was surrounded by arrows pointing in the same direction (→→→→→). On incongruent trials participants were asked to identify the direction of the center arrow which was surrounded by arrows pointing in the opposite direction (← ← → ← ←). The dependent variable was the mean response time on incongruent trials minus the mean response time on congruent trials.

Antisaccade. In the antisaccade task (Kane, Bleckley, Conway, & Engle, 2001), participants were asked to fixate on the center of the screen. A flashing '=' was presented on either side of the screen which participants needed to move their eyes toward (prosaccade condition) or away (antisaccade condition) from in order to identify a target letter presented briefly in the same (prosaccade) or opposite (antisaccade) position on the screen. A mask quickly followed the presentation of the target letter, so participants needed to engage attention control in the antisaccade condition to resist the prepotent tendency to look towards a flashing cue and instead look to the opposite side of the display where the target letter appeared. The dependent variable was target identification errors in the antisaccade condition of this task.

General fluid intelligence. gF was estimated from dependent variables from the RAPM, number series, and letter sets tasks. The individual tasks are described below and the estimate of gF is represented by the fluid intelligence latent variable in *Figure* 5.

RAPM. The traditional version of the RAPM task was used in the study conducted as part of the NSF grant. Thus, none of the alterations previously described in Experiments 1 and 2 were made to the RAPM task. Additionally, this version of the RAPM task imposed a ten-minute time limit on answering questions. As in Experiments 1 and 2, the dependent variable was the proportion of RAPM problems answered correctly.

Number series. In the number series task (Thurstone & Thurstone, 1962; Brewer & Unsworth, 2012), participants were shown a series of numbers and asked to determine the unstated rule among them. The solution to each problem was selected from a set of five possible solutions. Participants were given 4.5 minutes to complete as many of 15 problems as possible, and the dependent variable was the proportion of Number Series problems answered correctly.

Letter sets. In the letter sets task (Brewer & Unsworth, 2012), participants were shown five sets of letters containing four letters each and asked to determine the unstated rule shared by four of the five sets. The set of letters that did not follow an unstated rule shared by the other four sets was the correct solution on each problem in this task. Participants were given five minutes to complete as many of 20 problems as possible, and the dependent variable was the proportion of Letter Sets problems answered correctly.

Post-experimental questionnaire. In the post-experimental questionnaire, participants were asked to answer two constructive matching and two response elimination questions based on retrospective reports of strategic behavior on each of the three fluid intelligence tasks[13]. The dependent variables were the average response to the constructive matching and the average response to the response elimination questions for each of the three intelligence tasks. Constructive matching was estimated from the average constructive matching reported across all three of the intelligence tasks. Response elimination was estimated from the average response elimination reported across all three of the intelligence tasks. The estimates of constructive matching and response elimination are represented by the constructive matching and response elimination latent variables in *Figure* 5.

---

[13] The questions for number series and letter sets were modified from the RAPM strategic behavior questions to be relevant for the possible solutions in each task (see *Table* A in the appendix).

CHAPTER 10

EXPERIMENT 3 RESULTS

In Experiment 3, the variance shared across multiple measures of each construct was estimated and the models presented in *Figures* 3, 4, and 5 were fit to the data to assess H5 and H6. According to H5, cognitive processes will mediate the relation between WMC and gF, and H6 states that individual differences in strategic behavior will not mediate the relation between WMC and gF after accounting for individual differences in cognitive processes. In general, these hypotheses reflect the predictions made by Gonthier and Thomassin (2015). According to Gonthier and Thomassin (2015), individual differences in WMC are related to gF due to individual differences in cognitive processes, and these cognitive processes lead to differences in strategic behavior in the RAPM task. Descriptive statistics are presented for each of the variables in *Table* 8, and the correlations between the measures appear in *Table* 9.

Table 8

Descriptive Statistics for Experiment 3

Variable	M (SD)
Operation Span Partial-Unit Span Score	38.84 (8.04)
Symmetry Span Partial-Unit Span Score	19.85 (5.20)
Reading Span Partial-Unit Span Score	36.72 (8.84)
Color Arrays k	3.99 (1.03)
Orientation Arrays k	2.96 (1.20)
Immediate Free Recall - Primary Memory Estimate	8.79 (4.70)
Immediate Free Recall - Secondary Memory Estimate	20.18 (6.49)
Source Memory for Pictures Accuracy	0.82 (0.12)
Cued Paired Associates Total	14.10 (7.08)
Psychomotor Vigilance RT	521.11 (173.05)
Flanker RT Difference	73.30 (50.38)
Antisaccade Errors	0.35 (0.17)
RAPM Accuracy	0.50 (0.19)
Number Series Accuracy	0.60 (0.18)
Letter Sets Accuracy	0.52 (0.16)
RAPM Constructive Matching Average (Q1/Q2)	7.14 (1.58)
Number Series Constructive Matching Average (Q1/Q2)	7.77 (1.37)
Letter Sets Constructive Matching Average (Q1/Q2)	6.56 (1.76)
RAPM Response Elimination Average (Q1/Q2)	5.57 (1.99)
Number Series Response Elimination Average (Q1/Q2)	5.00 (2.48)
Letter Sets Response Elimination Average (Q1/Q2)	6.57 (1.85)

Table 9

Correlations for All Variables Modeled in Experiment 3

Variable No.	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	Operation Span Partial-Unit Span Score	--																				
2	Symmetry Span Partial-Unit Span Score	**0.39**	--																			
3	Reading Span Partial-Unit Span Score	**0.52**	**0.33**	--																		
4	Color Arrays k	**0.15**	**0.22**	**0.21**	--																	
5	Orientation Arrays k	**0.22**	**0.28**	**0.24**	**0.38**	--																
6	Immediate Free Recall - Primary Memory Estimate	0.01	**0.07**	0.04	**0.09**	0.07	--															
7	Immediate Free Recall - Secondary Memory Estimate	**0.21**	**0.14**	**0.29**	**0.16**	**0.20**	**-0.30**	--														
8	Source Memory for Pictures Accuracy	**0.10**	**0.17**	**0.15**	**0.20**	**0.27**	**0.19**	**0.20**	--													
9	Cued Paired Associates Total	**0.09**	**0.10**	**0.19**	**0.11**	**0.17**	**0.17**	**0.48**	**0.34**	--												
10	Psychomotor Vigilance RT	-0.06	-0.06	**-0.11**	**-0.16**	**-0.19**	-0.03	**-0.11**	**-0.15**	**-0.12**	--											
11	Flanker RT Difference	**-0.08**	**-0.12**	-0.02	-0.06	**-0.10**	**-0.09**	-0.06	**-0.09**	**-0.13**	**0.16**	--										
12	Antisaccade Errors	**-0.19**	**-0.28**	**-0.22**	**-0.24**	**-0.30**	**-0.09**	**-0.26**	**-0.26**	**-0.21**	**0.30**	**0.15**	--									
13	RAPM Total	**0.25**	**0.29**	**0.21**	**0.24**	**0.36**	**0.18**	**0.25**	**0.30**	**0.27**	**-0.19**	**-0.19**	**-0.39**	--								
14	Number Series Total	**0.31**	**0.28**	**0.27**	**0.21**	**0.29**	0.02	**0.23**	**0.12**	**0.14**	**-0.12**	**-0.15**	**-0.35**	**0.42**	--							
15	Letter Sets Total	**0.27**	**0.25**	**0.26**	**0.17**	**0.26**	**0.13**	**0.28**	**0.22**	**0.22**	**-0.16**	**-0.14**	**-0.34**	**0.39**	**0.42**	--						
16	RAPM Constructive Matching Average (Q1/Q2)	**0.14**	0.06	**0.15**	**0.16**	**0.14**	**0.08**	**0.09**	**0.13**	0.06	**-0.12**	-0.04	**-0.22**	**0.38**	**0.19**	**0.18**	--					
17	Number Series Constructive Matching Average (Q1/Q2)	**0.09**	0.05	**0.07**	**0.17**	**0.17**	0.05	**0.08**	**0.13**	0.07	**-0.14**	**-0.08**	**-0.13**	**0.17**	**0.22**	**0.18**	**0.45**	--				
18	Letter Sets Constructive Matching Average (Q1/Q2)	-0.03	-0.02	**0.07**	0.02	-0.02	0.02	**-0.07**	0.01	-0.06	**-0.07**	0.00	-0.03	0.04	-0.01	0.04	**0.27**	**0.29**	--			
19	RAPM Response Elimination Average (Q1/Q2)	-0.02	**-0.09**	0.00	-0.02	-0.07	0.00	**-0.16**	-0.02	**-0.14**	-0.05	0.06	0.05	-0.07	**-0.10**	**-0.08**	**0.18**	**0.16**	**0.29**	--		
20	Number Series Response Elimination Average (Q1/Q2)	**-0.07**	**-0.08**	-0.05	**-0.10**	**-0.14**	**-0.13**	**-0.19**	**-0.18**	**-0.23**	0.03	**0.10**	**0.19**	**-0.29**	**-0.18**	**-0.19**	0.03	**0.10**	**0.35**	**0.51**	--	
21	Letter Sets Response Elimination Average (Q1/Q2)	-0.05	0.00	0.03	0.05	0.04	0.01	0.06	0.06	0.02	-0.06	-0.02	-0.06	0.04	-0.01	**0.10**	**0.17**	**0.25**	**0.27**	**0.31**	**0.30**	--

Note: Signficiant correlations are in bold. Colors are discussed in the text.

There are several interesting things to note about the correlations in *Table* 9. With the exception of the primary memory estimate from the immediate free recall task, all of the tasks were correlated with other tasks within a factor (correlations in yellow). Consistent with much literature (Ackerman et al., 2005; Kane, Hambrick, & Conway, 2005), the WMC variables were positively correlated with the gF variables (correlations in green). WMC variables were only weakly correlated with constructive matching and response elimination strategic behavior (correlations in blue). The correlations between the WMC variables and the constructive matching variables were positive when significant relations were observed. Additionally, the correlations between the WMC variables and the response elimination variables were negative when significant relations were observed.

This is consistent with the predictions of Gonthier and Thomassin (2015), and because the individual complex span tasks contain task-specific variance no conclusions may be drawn about the source of these low correlations. A latent variable representing WMC may reveal stronger and consistent relations with the strategic behavior measures. However, rather than the issue being with the WMC variables, the lack of a strong relation between WMC and the strategic behavior variables is likely due to the positive correlations between the constructive matching variables and response elimination variables (correlations in orange). As in Experiments 1 and 2, this is likely due to a tendency to report using both strategies as a result of intraindividual variability in the application of one strategy versus the other on problems of increasing difficulty. This likely added unexplained variance to the measures of constructive matching and response elimination, which can be accounted for in a mediation model but likely masks relations when examining a correlation table.

Gonthier and Thomassin (2015) posited that individual differences in cognitive processes underlying WMC determine strategic behavior which influences gF task performance. When significant, the primary memory capacity and cue-dependent retrieval from secondary memory variables were positively correlated with the constructive matching variables (except the letter sets constructive matching variable which was negatively correlated with the secondary memory estimate from the immediate free recall task), and negatively correlated with the response elimination variables (correlations in purple). When significant, the attention control variables were negatively correlated with the constructive matching variables, and positively correlated with the response elimination variables (correlations in purple). These correlations are generally supportive of the theoretical framework proposed by Gonthier and Thomassin (2015), though the positive bias present in the measures of strategic behavior makes it difficult to fully evaluate support without fitting a model to the data that controls for this bias.

Generally, the gF variables were positively correlated with the constructive matching, and negatively correlated with the response elimination strategic behavior variables, consistent with Gonthier and Thomassin (2015), Bethell-Fox et al. (1984), and Vigneau et al (2006). However, the letter sets constructive matching variable was not related to performance on any of the gF tasks, and the only significant correlation between the gF tasks and the letter sets response elimination variable was positive (correlations in pink). Gonthier and Thomassin (2015) warned that the theoretical model they proposed is only applicable at present to RAPM task performance. Although the constructive matching and response elimination strategies are applicable to multiple choice tasks generally (Vigneau et al., 2006), this does not mean that these strategies are associated with better performance in other gF tasks the same way that they are in the RAPM task (Gonthier & Thomassin, 2015). Although the correlations between the strategic behavior variables and the other gF variables might indicate that strategic behavior varies across gF tasks in the present study, it is also possible that the positive bias in the measures of strategic behavior is once again masking the true effect.

As has been previously demonstrated by Unsworth et al. (2014), the working memory capacity complex span variables were positively correlated with performance on the primary memory capacity variables (except the immediate free recall primary memory estimate), positively correlated with the variables measuring cue-dependent retrieval from secondary memory, and negatively correlated with the attention control variables (correlations in gray). Additionally, the 1) primary memory capacity variables and the 2) cue-dependent retrieval from secondary memory variables were negatively correlated with the attention control variables, and positively correlated with the gF variables as well as each other (with the exception of the relation between the primary memory and secondary memory estimates from the immediate free recall task, which were negatively correlated). The attention control variables were negatively correlated with the gF variables due to the nature of the attention control tasks (correlations in gray). Slower RTs indicate poor attention control, as do more errors. By contrast, higher accuracy on the gF tasks indicates higher gF. As a result, the relation between attention control and gF variables should be negative rather than positive. The shared variance across each task within a construct (see *Table* 7) was estimated in models to assess the theoretical claim made by Gonthier and Thomassin (2015) that individual differences in cognitive processes that underlie WMC drive strategic behavior in gF tasks which influences performance on gF tasks (also see Unsworth & Redick, 2017).

. Anderson and Gerbing (1988) recommended first fitting a measurement model to the data and then fitting a structural model. According to the logic of this approach, the measurement model contains all possible associations among the latent variables which allows misfit in the structural model to represent paths omitted in the structural model. A measurement model was fit to the data including all possible correlations among the latent variables and then a structural model was fit to the data to elucidate the mechanisms underlying the relation between WMC and gF. Although the approach advocated by Anderson and Gerbing (1988; fitting a measurement model followed by a structural model) has not been taken in Experiments 1 and 2, this is due to the presence of a positive bias in the strategy questions due to their aggregate nature. To account for this bias, a correlation between the two strategic behavior latent variables was estimated and there were no paths omitted in the structural model (fit should be identical to the measurement model).

This is illustrated in *Table* 10 for the measurement and structural models fit to the data excluding cognitive processes. The data including only the strategic behavior variables as mediators were first submitted to a CFA in MPLUS (Muthén & Muthén, 1998-2011) with all possible associations among the latent variables, and the resulting measurement model is presented in *Figure* 13. Each of the items had moderately strong loadings on their latent factor, and all of the loadings across items within a factor are fairly comparable, indicating that the model was able to estimate shared variance across measures within a latent construct. The correlation between the two strategy factors was significant ($p < .001$) and positive, likely due to a slight bias to respond positively on all strategy questions due to their aggregate nature in the presence of intraindividual differences in strategic behavior across problems of increasing difficulty.

Table 10

Fit Indices for Strategic Behavior Measurement and Structural Models in Experiment 3

Model	χ^2	*df*	RMSEA	SRMR	CFI	TLI	AIC	BIC
Measurement	324.52	48.00	0.09	0.07	0.86	0.81	46387.10	46584.10
Structural	324.52	48.00	0.09	0.07	0.86	0.81	46387.10	46584.10

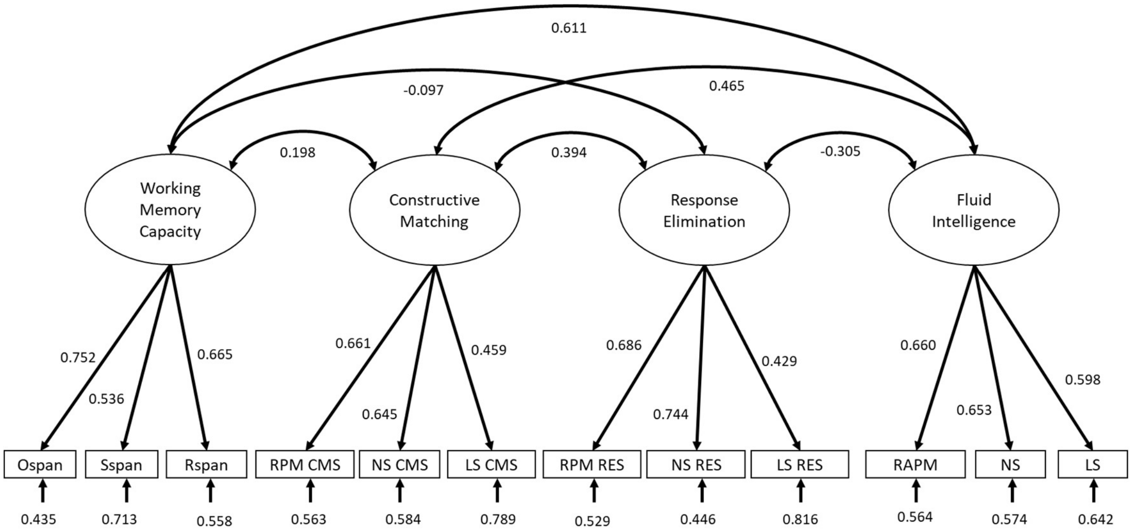


Figure 13. Measurement model for Experiment 3 strategic behavior data. Note: Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

The fit of the model in *Figure* 13 was just outside of an acceptable range of values, χ^2 (48) = 324.524, *p* < .01. An examination of the fit indices in *Table* 10 illustrates that the model was able to recreate the sample correlation matrix a little over 91% of the time (RMSEA and SRMR < .09). Additionally, both the CFI and TLI indicated that the model provided > 80% improvement over a baseline model that fixes the correlations across latent factors to 0. Next, a structural model was fit to the data and appears in *Figure* 14. The fit of this model is identical to the fit of the measurement model because no paths can be left out due to the positive bias present in the aggregate strategic behavior variables. The relation between WMC and RAPM task performance via constructive matching (standardized 95% bootstrapped CI[0.053, 0.171]), and via response elimination (standardized 95% bootstrapped CI[0.000, 0.094]) was significant. Thus, individual differences in strategic behavior mediated the relation between WMC and gF. Despite these mediating effects, a model excluding the direct path from WMC to fluid intelligence fit the data significantly worse than the model that freely estimated the regression of WMC on gF, $\Delta \chi^2 (1) = 61.646$, $p < .01$. WMC still predicted significant variance in gF ($p < .001$), indicating that additional processes also underlie the relation between WMC and gF (but see Rucker et al., 2011).

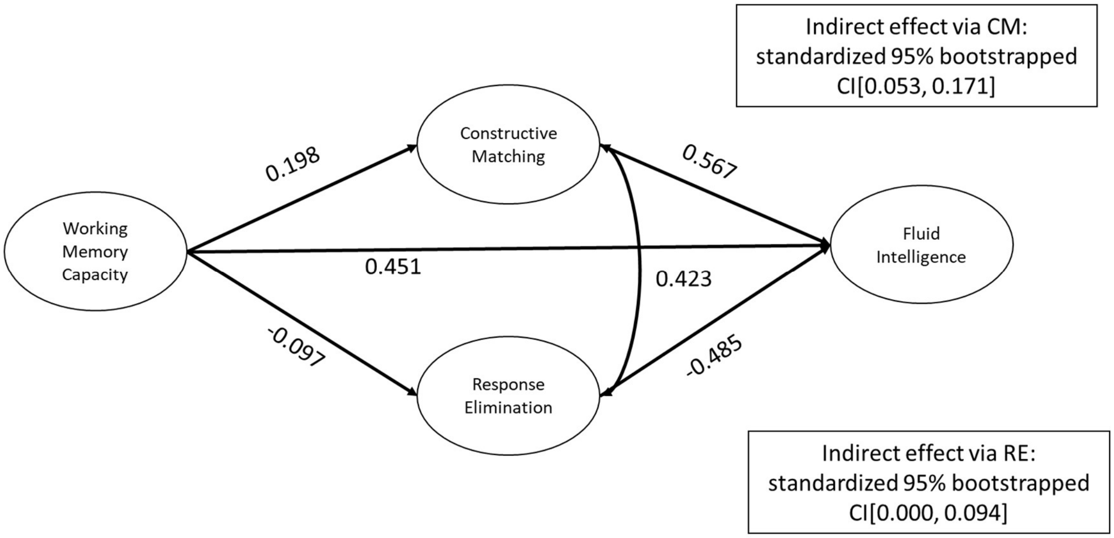


Figure 14. Structural model for Experiment 3 strategic behavior data. Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

The bias to respond positively was controlled for by allowing the strategy factors to correlate. In Experiments 1 and 2, the amount of unexplained variance this tendency to report using both strategies added reduced a signal that may have been otherwise present which likely resulted in reduced power to detect an effect. As shown in *Figure* 14, WMC predicted both constructive matching ($p < .001$) and response elimination ($p = .052$) strategic behavior. Additionally, the direction of the regression coefficients is consistent with the predictions in the literature. Participants with high working memory capacity were more likely to use a constructive matching strategy and participants with low working memory capacity were more likely to use a response elimination strategy (also see Jarosz & Wiley, 2012). The model in *Figure* 14 also indicates that the use of a constructive matching strategy was positively related to gF ($p < .001$) and the use of a response elimination strategy was negatively related to gF ($p = .001$). These effects are exactly what would be predicted based on a review of the literature (Bethel-Fox et al., 1984; Vigneau et al., 2006; Gonthier & Thomassin, 2015) and are consistent with the predictions in H1 and H3. Next, measurement and structural models including only the cognitive processes described in Unsworth et al. (2014) as mediators were fit to the data without these measures of strategic behavior.

The data including only the cognitive processes as mediators were first submitted to a CFA in MPLUS (Muthén & Muthén, 1998-2011) with all possible associations among the latent variables, and the resulting measurement model is presented in *Figure* 15. Each of the items had significant factor loadings. Although some loadings within a factor were stronger than other loadings for the primary memory capacity and attention control factors, these latent factors were comprised of more dissimilar tasks than the other latent factors so some variability in the size of the factor loadings within these factors may be expected. The correlation between primary memory capacity and cue-dependent retrieval from secondary memory was significant ($p < .001$), and both of these latent factors were positively related with each other and negatively related to attention control (*p*s < .001).

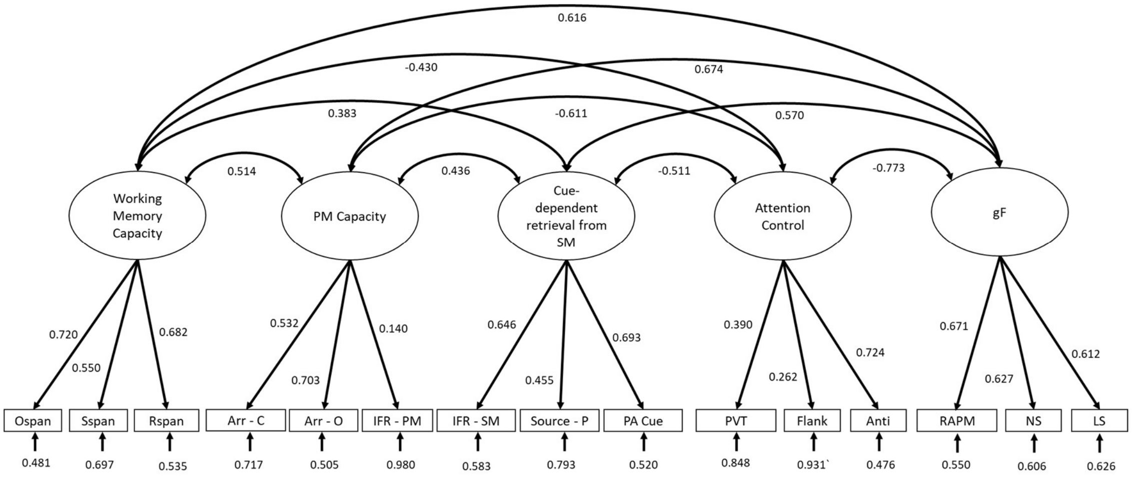


Figure 15. Measurement model 1 for Experiment 3 cognitive processes data. Note: Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

The presence of a correlation among strategy factors was previously used to justify correlating the strategy factors in the structural model to control for bias in the measures. The reason that approach was chosen was to control for bias variance that likely has nothing to do with constructive matching or response elimination strategic behavior (and the potential for this bias follows from Vigneau et al., 2006). The presence of correlations among the cognitive processes is not representative of any such biases. While constructive matching and response elimination strategic behavior should be negatively correlated in the absence of bias (as demonstrated in Jastrzębski et al., 2018), the correlations among the cognitive processes are likely more representative of meaningful effects in the literature rather than nonrandom error. Thus, these correlations will not be estimated in the structural model because the relations among these variables should be accounted for by their shared relation to WMC, gF, and the relation between WMC and gF.

The fit of the model in *Figure* 15 was outside of an acceptable range of values, χ^2 (80) = 498.73, $p < .01$. An examination of the fit indices in *Table* 11 illustrates that the model was able to recreate the sample correlation matrix a little over 91% of the time (RMSEA and SRMR < .09). The CFI and TLI indicated that the model provided > 76% improvement over a baseline model that fixes the correlations across latent factors to 0. These fit indices were less than ideal so the residuals were examined for the source of the misfit. The immediate free recall primary memory variable appeared to be a major source of misfit, and this variable also had the lowest factor loading in *Figure* 15. A second measurement model was fit to the data without that variable (the immediate free recall primary memory variable was dropped from the primary memory capacity factor) and the resulting model is presented *Figure* 16. The fit of this model was far more acceptable, χ^2 (67) = 220.871, $p < .01$. An examination of the fit indices in *Table* 11 illustrates that the second measurement model was able to recreate the sample correlation matrix a little over 94% of the time (RMSEA and SRMR < .06). The CFI and TLI indicated that the model provided at least a 90% improvement over a baseline model that fixes the correlations across latent factors to 0. The interpretation of measurement model 2 is the same as it was in measurement model 1.

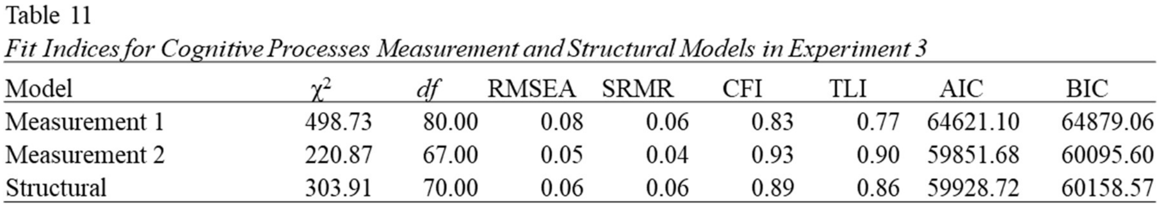


Table 11

Fit Indices for Cognitive Processes Measurement and Structural Models in Experiment 3

Model	χ^2	*df*	RMSEA	SRMR	CFI	TLI	AIC	BIC
Measurement 1	498.73	80.00	0.08	0.06	0.83	0.77	64621.10	64879.06
Measurement 2	220.87	67.00	0.05	0.04	0.93	0.90	59851.68	60095.60
Structural	303.91	70.00	0.06	0.06	0.89	0.86	59928.72	60158.57

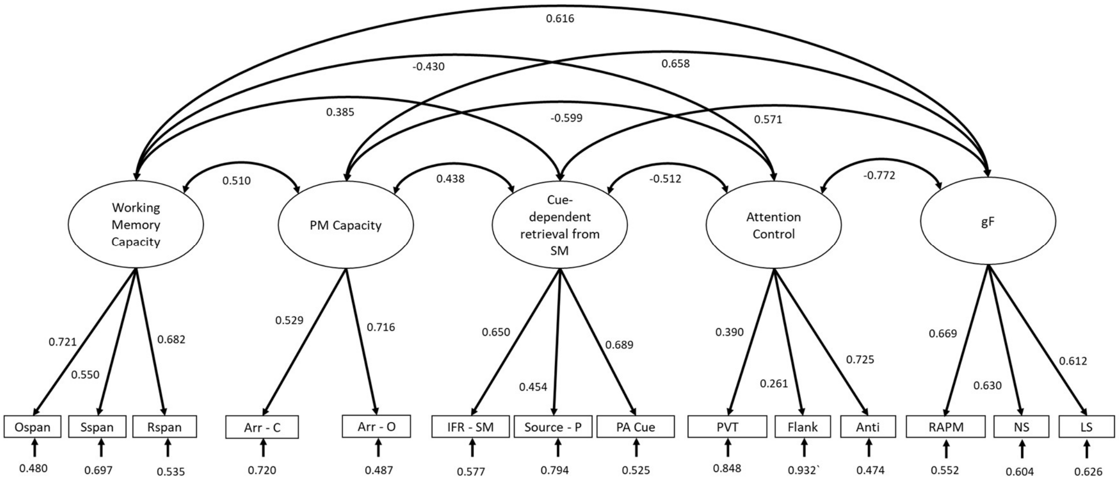


Figure 16. Measurement model 2 for Experiment 3 cognitive processes data. Note: Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

Next, a structural model was fit to the data (excluding the immediate free recall primary memory variable) and appears in *Figure* 17. This model fit the data significantly worse than the second measurement model, $\Delta \chi^2$ (3) = 83.038, $p < .01$, indicating that primary memory capacity, cue dependent retrieval from secondary memory, and attention control were correlated for reasons outside of the model. The relation between WMC and gF via primary memory capacity (standardized 95% bootstrapped CI[0.068, 0.249]), via cue-dependent retrieval from secondary memory (standardized 95% bootstrapped CI[0.050, 0.157]), and via attention control (standardized 95% bootstrapped CI[0.155, 0.356]) was significant. Thus, individual differences in cognitive processes mediated the relation between WMC and gF. A model excluding the direct path from WMC to fluid intelligence fit the data significantly worse than the model that freely estimated the regression of WMC on gF, $\Delta \chi^2$ (1) = 4.156, $p = .04$. WMC still predicted gF ($p = .031$), indicating that some other process also underlies the relation between WMC and gF (but see Rucker et al., 2011).

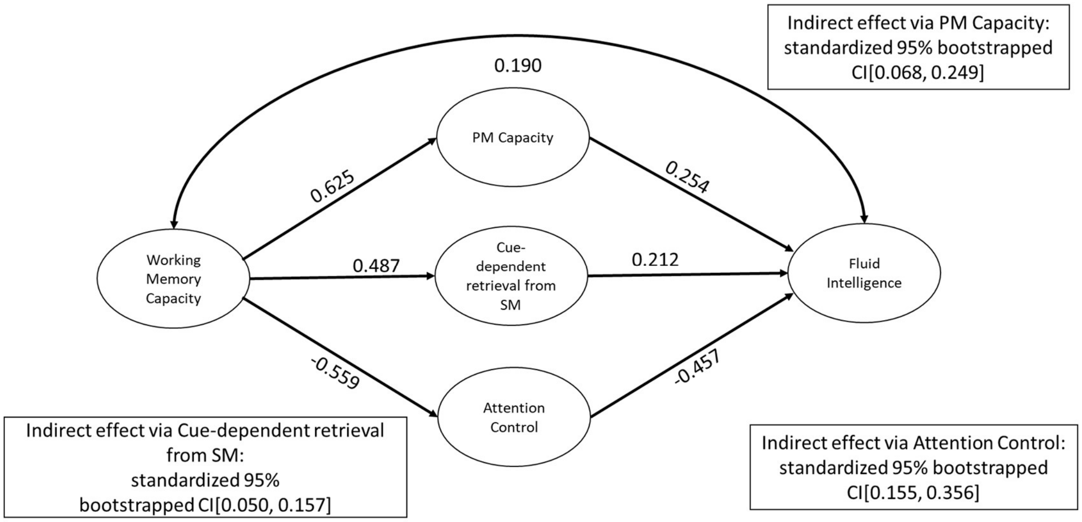


Figure 17. Structural model for Experiment 3 cognitive processes data. Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

The indirect effects via all three of the cognitive processes were significant. This is consistent with Unsworth et al. (2014), and in fact according to Rucker et al. (2011) researchers should not examine the direct effect to determine full vs. partial mediation (following recommendations in Baron & Kenny, 1986) because doing so can lead a researcher to conclude their mechanism(s) fully mediate some relation when in fact other mechanisms may be present (also see Tormala et al., 2007). Unsworth et al. (2014) suggest the presence of other processes that may underlie this relation (also see Unsworth & Redick, 2017). As a result, the present study was able to replicate Unsworth et al. (2014) and supported H5, indicating that the cognitive processes in *Figure* 17 represent important mechanisms underlying the relation between WMC and gF. H6 states that individual differences in strategic behavior will not mediate the relation between WMC and gF after accounting for individual differences in cognitive processes. To assess H6, the data including all cognitive processes and measures of strategic behavior as mediators were first submitted to a CFA in MPLUS (Muthén & Muthén, 1998-2011) with all possible associations among the latent variables, and the resulting measurement model is presented in *Figure* 18.

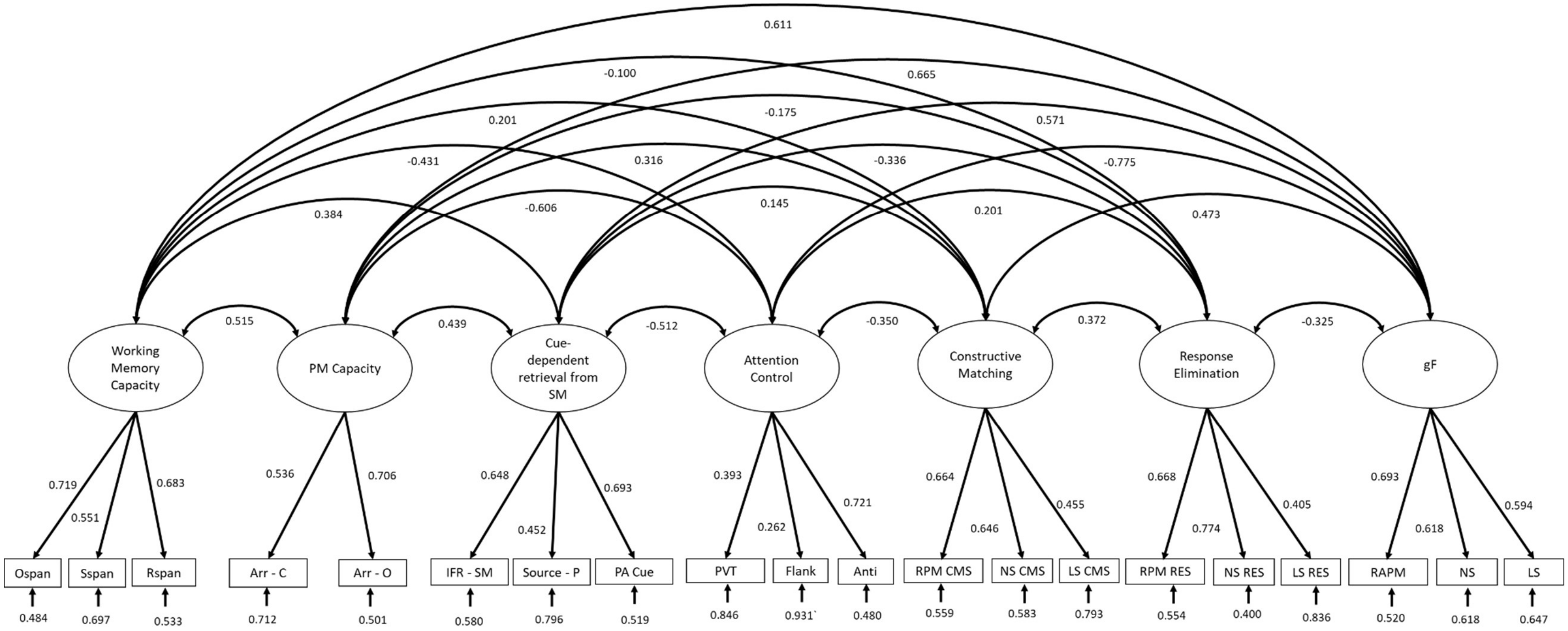


Figure 18. Measurement model with all mediators for Experiment 3 data. Note: Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

Each of the items had significant factor loadings. As before, some loadings within a factor were stronger than other loadings for the attention control factor, though this latent factor was comprised of more dissimilar tasks than the other latent factors so some variability in the size of the factor loadings within these factors may be expected. The two measures of strategic behavior were positively related ($p < .001$), once again likely due to the aggregate nature of the measures. Similar to Unsworth et al. (2014), WMC was positively related to primary memory capacity, cue-dependent retrieval from secondary memory, and gF, and WMC was negatively related to attention control. The correlation between primary memory capacity and cue-dependent retrieval from secondary memory was significant ($p < .001$), and both of these latent factors were positively related with each other and negatively related to attention control (*p*s < .001). Attention control was negatively related to gF, also consistent with Unsworth et al. (2014). This negative relation between the latent constructs in the model and attention control is a result of the choice of variables measuring attention control. Larger response times in Flanker and the PVT as well as more errors on the Antisaccade task are indicative of poor attention control resulting in the observed negative direction of the correlations.

Consistent with Gonthier and Thomassin (2015; also see Jarosz & Wiley, 2012 and Bethel-Fox et al., 1984), working memory capacity and gF were both positively correlated with constructive matching and negatively correlated with response elimination. The fit of the model in *Figure* 18 was acceptable though some of the fit indices were outside of an acceptable range, χ^2 (149) = 594.061, $p < .01$. An examination of the fit indices in *Table* 12 illustrates that the measurement model was able to recreate the sample correlation matrix a little over 93% of the time (RMSEA and SRMR < .07).

The CFI and TLI indicated that the model provided at least an 82% improvement over a baseline model that fixes the correlations across latent factors to 0. The positive bias likely present in the aggregate strategy measures can be better accounted for by allowing all the strategy variables to load onto a positive bias factor that is constrained to be uncorrelated with the factors representing theoretical constructs (the bias should not be related to the variables of interest).

Table 12

Fit Indices for Measurement and Structural Models with all Mediators in Experiment 3

Model	χ^2	*df*	RMSEA	SRMR	CFI	TLI	AIC	BIC
Measurement Model 1	594.06	149.00	0.06	0.06	0.87	0.83	78437.52	78817.47
Measurement Model 2	412.24	143.00	0.05	0.04	0.92	0.89	78267.70	78675.80
Structural Model 1	595.55	153.00	0.06	0.05	0.87	0.83	78431.00	78792.20
Structural Model 2	535.18	152.00	0.06	0.05	0.88	0.86	78372.63	78738.52

A second measurement model was fit to the data that included a positive bias factor, and the resulting model is presented in *Figure* 19. The fit of this model was significantly better than the fit of the first measurement model, $\Delta \chi^2$ (6) = 181.816, $p <$ .01. An examination of the fit indices in *Table* 12 illustrates that the second measurement model was able to recreate the sample correlation matrix a little over 95% of the time (RMSEA and SRMR < .05). The CFI and TLI indicated that the model provided at least an 89% improvement over a baseline model that fixes the correlations across latent factors to 0. Further, the AIC and BIC were lower in Measurement Model 2 compared to Measurement Model 1, indicating that the additional parameters improve model fit despite a penalty for model complexity. As a result, this measurement model was retained and used to compare with the fit of a structural model.

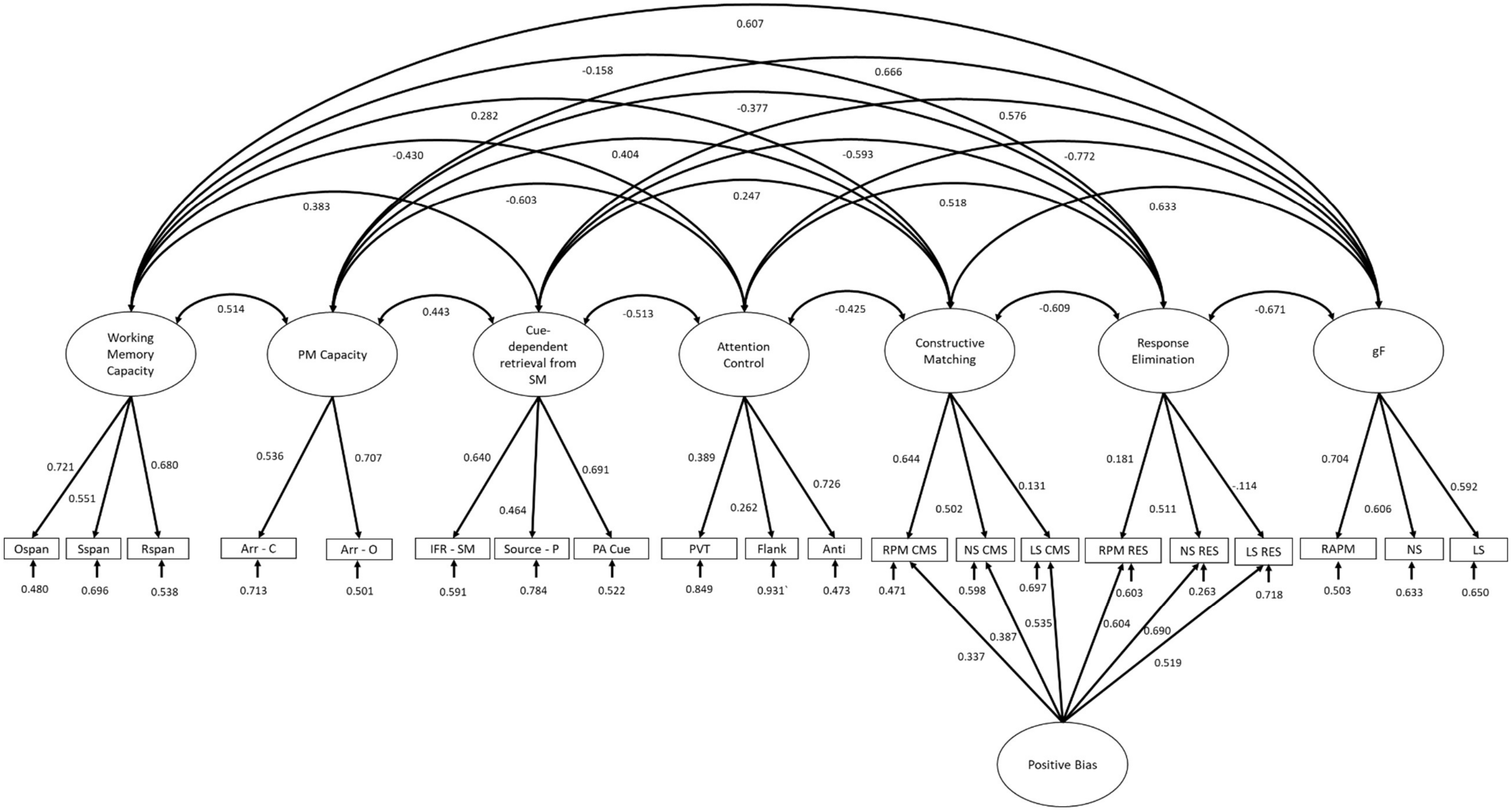


Figure 19. Measurement Model 2 with all mediators for Experiment 3 data. Note: Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

Only one factor loading changed direction in the second measurement model. People who reported more response elimination on the RAPM and number series tasks tended to report less response elimination on the letter sets task. Gonthier and Thomassin (2015) note that other measures of gF may not benefit from the same strategic behavior that benefits participants completing an RAPM task. Just because a strategy is applicable (Vigneau et al., 2006) does not speak to whether the strategy should be used during a specific task. Thus, although it is not ideal that the factor loading for response elimination strategic behavior in the letter sets task is negative ($p = .045$), it is not necessarily indicative of a problem. The factor loading for response elimination strategic behavior in the RAPM task was small (.181), but significant, $p = .001$. This reflects the cost of collecting data containing nonrandom error. This is the size of the signal when nonrandom error variance (positive bias due to aggregate measurement) is partitioned out. The factor loadings for the items in the positive bias factor were all moderately strong indicating that all of the strategy questions shared variance that was unrelated to the theoretical constructs of interest.

After controlling for this positive bias, the correlation between WMC and the two indices of strategic behavior became larger (without changing direction). WMC was positively related to the use of a constructive matching strategy and negatively related to the use of a response elimination strategy, consistent with Gonthier and Thomassin (2015) and Jarosz and Wiley (2012). The correlation estimates for the relation between WMC and each of the cognitive processes, and with gF, hardly changed at all in Measurement Model 2. Additionally, the correlations between each of the three cognitive processes and gF hardly changed in Measurement Model 2. The constructive matching factor was still positively related to primary memory capacity, cue-dependent retrieval from secondary memory, and gF. Additionally, the constructive matching factor was still negatively related to the attention control factor. However, in the absence of positive bias stronger relations were observed. The correlations between response elimination and each of the cognitive processes, and between response elimination and gF also did not change direction. Once again, removing the positive bias allowed stronger relations between the theoretical constructs to emerge.

Interestingly, after controlling for positive bias there was a large negative correlation between constructive matching and response elimination strategic behavior (consistent with Jastrzębski et al., 2018). This indicates that participants tended to use one strategy or the other on a task. Participants who used a constructive matching strategy were less likely to use a response elimination strategy, and participants who used a response elimination strategy were less likely to use a constructive matching strategy (indicating that it was a positive bias we removed, which is consistent with Bethell-Fox et al., 1984, Vigneau et al., 2006, and Gonthier and Thomassin, 2015). This negative relation was not observed in Experiment 1 or Experiment 2. This is because there was not enough information collected (known information) to fit a model with a positive bias factor. This combined with the larger sample size in Experiment 3 allowed for a fairer assessment of the theoretical stance taken by Gonthier and Thomassin (2015).

Next, a structural model was fit to the data (excluding the immediate free recall primary memory variable) and appears in *Figure* 20 (the positive bias factor is not presented because it was constrained to be uncorrelated with all the theoretical factors represented in the structural model). This model fit the data significantly worse than the second measurement model, $\Delta \chi^2$ (10) = 183.304, $p < .01$, indicating that there were correlations between the latent constructs for reasons outside of the model. The relation between WMC and gF via constructive matching (standardized 95% bootstrapped CI[0.058, 0.148]), via response elimination (standardized 95% bootstrapped CI[0.015, 0.144]), via primary memory capacity (standardized 95% bootstrapped CI[0.003, 0.211]), via cue-dependent retrieval from secondary memory (standardized 95% bootstrapped CI[0.006, 0.127]), and via attention control (standardized 95% bootstrapped CI[0.107, 0.321]) was significant. Thus, individual differences in cognitive processes and strategic behavior mediated the relation between WMC and gF. A model excluding the direct path from WMC to fluid intelligence fit the data significantly worse than the model that freely estimated the regression of WMC on gF, $\Delta \chi^2$ (1) = 3.718, $p = .05$. WMC still predicted gF ($p = .048$), indicating that some other process also underlies the relation between WMC and gF.

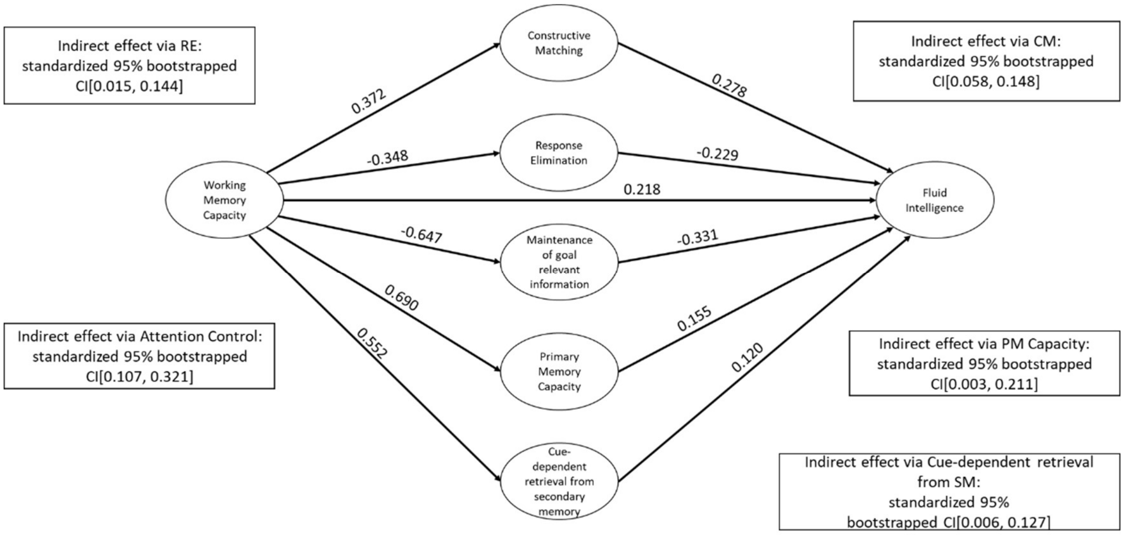


Figure 20. Structural model for Experiment 3 data with all mediators. Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

This model differed from the models containing only cognitive processes as mediators and the models only containing the strategic behavior measures as mediators in the way that the positive bias in the measures of strategic behavior was handled. As a result, the size of the indirect effects cannot be compared meaningfully to evaluate the predictions outlined in the introduction. Although the data indicate that collectively these mediators only partially mediate the relation between WMC and gF, Rucker et al. (2011) argued that researchers should not use the direct effect to make any claims about mediation. Additionally, Rucker et al. (2011) indicated that an independent variable (WMC) may be more related to a mediator than the dependent variable. As a result, the indirect effect may be stronger than a direct effect. Rucker et al. (2011) recommended focusing on the magnitude of the indirect effect when evaluating mediation models. This highlights an issue with the predictions in H6 because by extension it is plausible that an indirect effect of WMC via the cognitive processes via strategic behavior could similarly exert a stronger effect on RAPM task performance than the model presented in *Figure* 20.

A second structural model was fit to the data (excluding the immediate free recall primary memory variable) and appears in *Figure* 21 (the positive bias factor is not presented because it was constrained to be uncorrelated with all the theoretical factors represented in the structural model). This model reflects the theoretical stance taken by Gonthier and Thomassin (2015) that WMC and gF are related because individual differences in cognitive processes underlie WMC and determine strategic behavior which influences gF (also see Unsworth & Redick, 2017). This model fit the data significantly better than the previous structural model, $\Delta \chi^2$ (1) = 60.372, $p < .01$. However, the model still fit the data significantly worse than the second measurement model, $\Delta \chi^2$ (9) = 122.932, $p < .01$, indicating that there were relations present between the latent constructs for reasons outside of the model.

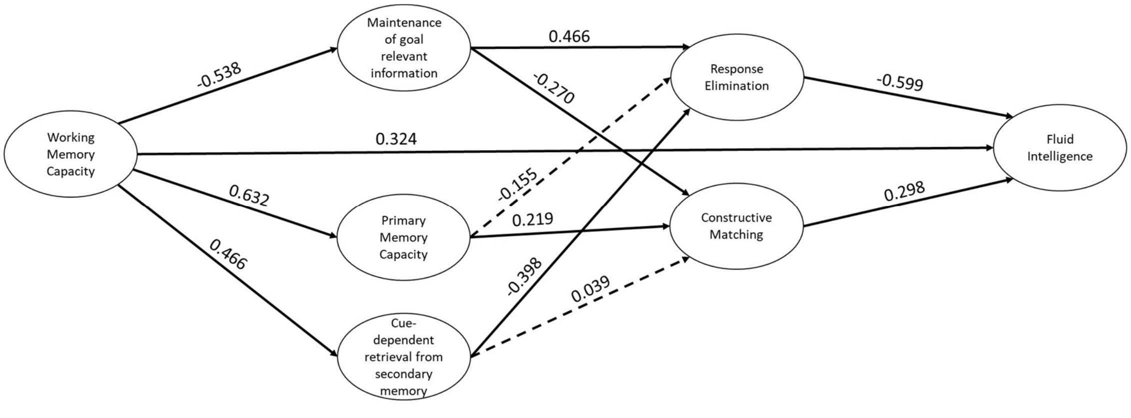


Figure 21. Structural model 2 for Experiment 3 data with all mediators, based on the theory outlined by Gonthier and Thomassin (2015). Latent factor variances are fixed to 1 and latent factor means are fixed to 0. Described in detail in the text.

The indirect effect of WMC on gF via attention control was significant via response elimination (standardized 95% bootstrapped CI[0.037, 0.263]), but not constructive matching (standardized 95% bootstrapped CI[-0.004, 0.091]). This indicates that part of the reason that WMC and gF share variance is because individual differences in attention control (goal maintenance) influence response elimination strategic behavior. In Experiment 3 the traditional version of the RAPM task was used so goals to consider the most appropriate strategy needed to be maintained during the same time as information specific to the RAPM problem (compared to Experiment 2). This particular type of goal maintenance masked a relation between WMC and RAPM task performance in Experiment 2. However, in Experiment 3 the task was not structured in a way that facilitated the separate maintenance of the goal to consider evidence regarding the effectiveness of strategic behavior. The results of exploratory analyses conducted in Experiment 2 indicated that accounting for this variance separately from other goal maintenance variance should allow more of the shared variance between WMC and gF to be estimated.

In Experiment 3, the indirect effect of WMC on gF via attention control was not significant via constructive matching. By contrast, the indirect effect of WMC on gF via attention control was significant via response elimination. In Experiment 2 low WMC participants used a constructive matching strategy more in the control condition, and there was no further benefit to providing a constructive matching manipulation that provided environmental support for the use of that strategy. This may result from the amount of information that low WMC participants can hold in primary memory. The increase in the performance of low WMC participants in the control condition in

Experiment 2 compared to Experiment 1 may have resulted from more information about problem specific goal maintenance being represented in primary memory after the need for maintenance of information about strategic behavior is minimized. Additionally, high WMC participants do not benefit from additional time to reflect on strategic behavior in the control condition for Experiment 2 compared to Experiment 1.

However, RAPM task performance for high WMC participants in the constructive matching condition in Experiment 2 was higher than high WMC participants in the constructive matching condition in Experiment 1. Thus, more information about the problem can be held in primary memory for high WMC participants as well indicating that allowing participants time to reflect on their strategies may be beneficial to low WMC and high WMC participants differentially. In Experiment 2, primary memory did not need to contain as much information about a participant's strategic behavior because the participant was asked to report strategic behavior after answering each question. In Experiment 3, strategic behavior is only assessed at the end of the three intelligence tasks. Thus, in Experiment 3 the goal to consider strategic behavior must be prioritized and stored in primary memory where it must be maintained until the end of the tasks. This may result in participants storing less problem specific information (due to the capacity of primary memory) resulting in worse RAPM task performance.

If the capacity of primary memory limits how much information may be stored in primary memory, the maintenance of goal-relevant information (attention control) is needed to determine what information gains access to primary memory. If information about the appropriate strategic behavior to use needs to be stored in primary memory in Experiment 3, then information about strategic behavior will be prioritized via attention control. This may result in a reduction of the overall capacity of primary memory reserved for problem specific goal maintenance and may reduce gF. Although the indirect effect from WMC to gF via attention control was only significant via response elimination strategic behavior, providing participants with more time to reflect on their strategic behavior appeared to increase the use of constructive matching strategic behavior. This global goal to evaluate whether the strategy being used is effective requires more maintenance when you are not provided extra time, but this information is not specific to the specific problem being solved so these data indicate that future research should collect measures of strategic behavior after each RAPM question (depending on the nature of the question being asked).

The indirect effect of WMC on gF via primary memory capacity was not significant via response elimination (standardized 95% bootstrapped CI[-0.046, 0.163]), nor via constructive matching (standardized 95% bootstrapped CI[-0.002, 0.085]). Although primary memory capacity should limit the amount of information that can be constructed for a given gF problem (in line with the significant indirect effect via constructive matching), primary memory capacity is not likely relevant for simple comparisons of each response option to the problem matrix. As indicated in the discussion of the Experiment 2 results, response elimination may rely more on cue-dependent retrieval of information from secondary memory than primary memory capacity.

The indirect effect of WMC on gF via cue-dependent retrieval from secondary memory was significant via response elimination (standardized 95% bootstrapped CI[0.057, 0.165]), but was not significant via constructive matching (standardized 95%

bootstrapped CI[-0.018, 0.029]). Again, the information relevant to assessing strategic behavior was reported after each RAPM problem in Experiment 2 which minimized the need to maintain the goal of considering strategic behavior. This goal in particular is important to consider because the correlation between WMC and gF was similar across strategy conditions in Experiment 2, though an examination of the underlying data indicate that providing time to consider strategic behavior had differential effects on strategic behavior and RAPM task performance for low WMC compared to high WMC participants across strategy conditions. Response elimination strategic behavior does not rely on the construction of a solution that takes all relevant information into consideration. Rather, it requires examining each potential solution for a match. Thus, response elimination strategic behavior should relate to cue-dependent retrieval from secondary memory.

Overall, there was support for the notion that individual differences in WMC relate to gF because cognitive processes lead to differences in strategic behavior which in turn affects how well participants perform on gF tasks. Although there was still a relation between WMC and gF ($p < .001$), Rucker et al. (2011) cautions against interpreting this presence of a direct effect (or its absence). In fact, 91.9% of the variance in gF was accounted for in the model presented in *Figure* 21. Variance partitioning was performed by conducting a series of regression analyses on the factor scores saved while fitting the model presented in *Figure* 21 to allocate the overall R^2 into shared and unique variance (see *Table* 13). The results of the variance partitioning were generally supportive of the model proposed by Gonthier and Thomassin (2015). The two measures of strategic behavior accounted for shared variance between WMC and gF via the three cognitive processes. However, these measures of strategic behavior also accounted for unique variance in gF, indicating that Vigneau et al. (2006) may have been correct to conclude that not all of the variance in gF attributable to strategic behavior overlaps with variance in WMC.

Table 13

R2 Values for Regression Analyses Predicting gF From Various Predictor Variables

Predictor Variables	R^2
WMC, PMC, CDSM, ATTN, CM, RE	0.996
WMC, CDSM, ATTN, CM, RE	0.995
WMC, PMC, ATTN, CM, RE	0.993
WMC, PMC, CDSM, CM, RE	0.994
WMC, PMC, CDSM, ATTN, RE	0.953
WMC, PMC, CDSM, ATTN, CM	0.924
PMC, CDSM, ATTN, CM, RE	0.967
WMC, ATTN, CM, RE	0.993
WMC, CDSM, CM, RE	0.994
WMC, CDSM, ATTN, RE	0.953
WMC, CDSM, ATTN, CM	0.923
CDSM, ATTN, CM, RE	0.955
WMC, PMC, CM, RE	0.992
WMC, PMC, ATTN, RE	0.948
WMC, PMC, ATTN, CM	0.890
WMC, PMC, CDSM, RE	0.953
WMC, PMC, CDSM, CM	0.887
WMC, PMC, CDSM, ATTN	0.855
PMC, ATTN, CM, RE	0.967
PMC, CDSM, CM, RE	0.967
PMC, CDSM, ATTN, RE	0.930
PMC, CDSM, ATTN, CM	0.901
WMC, CM, RE	0.992
WMC, ATTN, RE	0.947
WMC, ATTN, CM	0.888
ATTN, CM, RE	0.955
WMC, CDSM, RE	0.952
WMC, CDSM, CM	0.885
CDSM, CM, RE	0.955
WMC, CDSM, ATTN	0.841
CDSM, ATTN, RE	0.909
CDSM, ATTN, CM	0.870
WMC, PMC, RE	0.948
WMC, PMC, CM	0.845
PMC, CM, RE	0.967
WMC, PMC, ATTN	0.815
PMC, ATTN, RE	0.929
PMC, ATTN, CM	0.839
WMC, PMC, CDSM	0.765
PMC, CDSM, RE	0.929
PMC, CDSM, CM	0.825
PMC, CDSM, ATTN	0.838
WMC, CM	0.843
WMC, RE	0.947
CM, RE	0.955
WMC, ATT	0.802
ATTN, RE	0.908
ATTN, CM	0.773
WMC, CDSM	0.744
CDSM, RE	0.907
CDSM, CM	0.742
CDSM, ATTN	0.776
WMC, PMC	0.710
PMC, RE	0.927
PMC, CM	0.845
PMC, ATTN	0.774
PMC, CDSM	0.695
WMC	0.688
PMC	0.556
CDSM	0.471
ATTN	0.661
CM	0.505
RE	0.905

CHAPTER 11

GENERAL DISCUSSION

The present study aimed to provide a mechanistic account of the relation between WMC and gF (e.g., see Unsworth and Redick, 2017). The goal of the first two experiments was to replicate the results of Gonthier and Thomassin (2015) showing that individual differences in strategic behavior mediate the relation between WMC and gF (H1/H3), and that experimentally manipulating strategic behavior in an RAPM task reduces the relation between WMC and gF (H2/H4). In Experiment 1, there was a selective enhancement in RAPM task performance for low working memory capacity participants in the constructive matching condition. Thus, there was support for H2 in Experiment 1. However, there was no support for H1 in Experiment 1. The results of Experiment 1 may indicate that we were unable to replicate Gonthier and Thomassin (2015). However, the measures of strategic behavior may not have been sensitive to the experimental question and may have reduced the ability to detect an effect if there is one. In Experiments 1-3, the retrospective measures of strategic behavior were positively correlated indicating that more reported use of one strategy was associated with more reported use of the other strategy. This positive correlation likely reflects the use of aggregate measures of strategic behavior when intraindividual variability in strategic behavior is expected (Vigneau et al., 2006).

Although this positive bias may have hindered the ability to support H1 in Experiment 1, H2 does not make any predictions based upon self-report measures. In Experiment 1, a selective enhancement in RAPM task performance was observed for low working memory capacity participants in the constructive matching condition. The correlation between WMC and RAPM task performance was not significant in the constructive matching condition (but was significant in a control condition), indicating that differences in the application of a constructive matching strategy underlie the shared variance between WMC and RAPM task performance in a more traditional version of the task. Thus, the failure to support H1 and replicate Gonthier and Thomassin (2015) in Experiment 1 may have been due to either A) an additional failed replication of Gonthier and Thomassin (2015; Jastrzębski et al., 2018), or B) the contamination of the self-report measures with nonrandom error that reduced power to detect an effect. As a result, in Experiment 2 measures of strategic behavior were collected after each RAPM problem and a measure of intraindividual variability in strategic behavior was estimated for each participant.

The goal of Experiment 2 was to develop more sensitive measures of strategic behavior that did not contain this positive bias. However, asking participants to report strategic behavior after each RAPM problem appeared to change the nature of the relation between WMC and gF (see *Figures* 8 and 10). The relation between WMC and gF was unchanged across strategy conditions in Experiment 2 (there was no support for H4). However, low WMC participants performed better in the control condition in Experiment 2 compared to Experiment 1, and high WMC participants performed better in the constructive matching condition in Experiment 2 compared to Experiment 1. The data in Experiment 2 did not support H4 (H2), indicating that the reason individual differences in constructive matching strategic behavior underlie the relation between WMC and gF (support for H2 was found in Experiment 1) is due to reflection on the results of different types of strategic behavior.

However, frequent reports of strategic behavior may have minimized the need for goal maintenance processes related to this information and reduced the amount of information being held in primary memory allowing participants to maintain more information relevant to solving the problem. The data in Experiment 2 did not support H3. Even after removing variance in strategic behavior not related to solving the specific RAPM problem, individual differences in strategic behavior on an RAPM task did not mediate the relation between WMC and gF. The opposing influences of a manipulation of strategic behavior (which should be associated with a reduced correlation between WMC and gF) and a removal of goal maintenance variance unrelated to the specific RAPM problem (which should be associated with an increase in the correlation between WMC and gF) resulted in no overall change in the correlation between WMC and gF across strategy conditions.

The shared variance between WMC and gF was further examined in Experiment 3 by fitting a model to data that included measures of all of the cognitive processes as well as measures of strategic behavior. Experiment 3 power concerns (due to the positive bias present in the aggregate measures of strategic behavior) were addressed with a larger sample size rather than addressing the precision of the measures as in Experiment 2. The df gained by including multiple measures of cognitive processes mediators as well allowed for a method factor to be fit to the strategic behavior data to capture the shared variance across all measures due to the positive bias in responding to using both strategies over the course of the entire task. A method factor representing positive bias was fit to the strategic behavior data for each of the gF tasks.

This positive bias factor was constrained to be uncorrelated with any of the constructs of interest in the model. This assumption follows from the idea that this bias arises due to an insensitivity of the measures to intraindividual variability that is thought to be present in the data rather than due to characteristics of the individual themselves. In the measurement model in Experiment 3 with a method factor, the correlation between constructive matching and response elimination strategic behavior (at the latent level) was negative as would be predicted if the measures of strategic behavior represented two alternative behaviors that participants might adopt (constructive matching or response elimination, Bethell-Fox et al., 1984; Vigneau et al., 2006). This negative correlation between the constructive matching and response elimination latent factors was in demonstrated Jastrzębski et al. (2018).

In Experiment 3 there was support for the idea put forth in Gonthier and Thomassin (2015) that individual differences in cognitive processes underlying WMC drive differences in strategic behavior which determine performance on a gF task (also see Unsworth and Redick, 2017). The indirect effects of WMC on gF via each of the cognitive processes and their contribution to constructive matching and response elimination strategic behavior presented conflicting support for the theory proposed in Gonthier and Thomassin (2015). However, there were relations in the data unaccounted for in the model presented in *Figure* 21, indicating that the theoretical stance taken by Gonthier and Thomassin (2015) needs to be revised to account for residual relations among latent constructs. However, 91.9% of the variance in gF was accounted for in the model presented in *Figure* 21, indicating that the model proposed by Gonthier and Thomassin (2015) represented an adequate explanation of the gF variance.

Unlike Experiments 1 and 2, and inconsistent with Jastrzębski et al. (2018), there was support for H1/H3 in Experiment 3. The models that did not include any of the cognitive processes did not contain enough known information to fit a positive bias factor. In these initial models the positive bias present in the retrospective reports of strategic behavior was accounted for by allowing the strategic behavior latent variables to correlate (similar to Experiments 1 and 2). In the model presented in *Figure* 14, the indirect effect from WMC to RAPM task performance via both constructive matching and response elimination strategic behavior was significant, indicating that individual differences in strategic behavior underlie the relation between WMC and RAPM task performance (consistent with Gonthier and Thomassin, 2015). The models that did not include any of the strategic behavior mediators provided support for H5, which stated that the cognitive processes would mediate the relation between WMC and gF (consistent with Unsworth et al., 2014).

Rucker et al. (2011) argued that all mediation is partial mediation and thus it is not meaningful to state that the models including only strategic behavior mediators or cognitive processes mediators only partially mediated the relation between WMC and gF. Following the recommendations of Rucker et al. (2011) to examine only the theorized indirect effects in mediation models, the model in *Figure* 21 (which contained the positive bias method factor) was fit to the data to assess the theoretical stance taken by Gonthier and Thomassin (2015) that individual differences in cognitive processes underlie WMC and drive differences in strategic behavior on gF tasks, which subsequently influences performance on the gF task. Examination of the indirect effects in the model in *Figure* 21 generally supported the predictions in Gonthier and Thomassin (2015).

Findings from the exploratory analyses support the idea that merely examining the correlation is not enough. The measure of strategic behavior assessed after each RAPM problem may have allowed for a more valid representation of information relevant for goal maintenance processes required for solving each problem and by minimizing the need to maintain the goal of reflecting on strategic behavior. These possible opposing effects critically highlight the possibility that there may be no changes in task performance, yet this does not necessarily mean that a process does not underlie a relation. It was only by measuring both constructive matching and control strategic behavior across Experiments that we were able to obtain some preliminary evidence. In Experiment 2 the overall size of the correlation was unchanged by the constructive matching manipulation.

However, support for H2/H4 was previously obtained in Experiment 1. Engle and Kane (2004) posited that individual differences in strategic behavior represent a nuisance variable in describing the relation between WMC and gF on the basis of evidence that the correlation between WMC and gF increases when differences in strategic behavior are controlled for. However, the results of Gonthier and Thomassin (2015) and the present study supported H2, indicating that the correlation between WMC and gF actually decreases significantly when differences in strategic behavior are minimized. This indicates that rather than representing a nuisance variable, individual differences in strategic behavior represent a critical mechanism underlying the relation between WMC and gF. Across three experiments, there was support for the notion that variability in gF arises due to individual differences in strategic behavior driven by variation in cognitive processes underlying WMC.

This directly counters the notion that individual differences in strategic behavior represent a nuisance variable in describing the relation between WMC and gF (Engle & Kane, 2004). One potential limitation in the present study is the assumption that the strategic behavior endorsed in an RAPM task should be the strategic behavior endorsed in the number series and letter sets tasks. In fact, Gonthier and Thomassin (2015) warn against making this assumption and data from Experiment 3 in the present study illustrate why this assumption may present an issue. The factor loading for letter sets response elimination strategic behavior was negative on the response elimination factor, while the other two measures of response elimination loaded positively onto the factor. This could reflect differences across the tasks. For both number series and the RAPM, participants were asked to select the correct solution from a set of solutions that followed some unstated rule or set of rules. For the letter sets task, participants were asked to select the solution that did not follow the same rule or set of rules as the others.

Thus, there is reason to believe that more reliance on response elimination should be expected in letter sets compared to number series or the RAPM. This highlights a potential concern in the interpretation of the results from the present study. If response elimination is an efficient strategy in letter sets, then the predictions about how response elimination relates to WMC and gF would be different for letter sets compared to number series and the RAPM task. Future research should evaluate the model proposed by Gonthier and Thomassin (2015) after giving further consideration to the measures of strategic behavior that reflect strategic behavior across the different gF tasks. However, the data from Experiment 3 supported the model proposed by Gonthier and Thomassin (2015) and are in agreement with Unsworth et al.'s (2014; also see Unsworth and Redick, 2017) statement that interactions between the cognitive processes proposed in their model and other potential mediators are important to consider in modeling the shared variance between WMC and gF. Additional research should assess the contribution of goal maintenance, primary memory capacity, and cue-dependent retrieval from secondary memory to strategic behavior that is assessed after each gF problem, and when it is assessed only once at the end of the task. These data will allow researchers to assess the contribution of each of the cognitive processes to strategic behavior.

REFERENCES

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological bulletin, 131*(1), 30-60. doi: 10.1037/0033-2909.131.1.30

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423. doi: 10.1037/0033-2909.103.3.411

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47-89. doi: 10.1016/S0079-7421(08)60452-1

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, *8*(3), 205-238. doi: 10.1016/0160-2896(84)90009-6

Bollen, K. A. (1989). *Structural Equations with Latent Variables.* New York: John Wiley & Sons.

Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language, 66*(3), 407-415. doi: 10.1016/j.jml.2011.12.009

Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review, 13*(5), 826-830.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action.* Boston: Houghton Mifflin.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41*(4), 244-262. doi: 10.1016/j.intell.2013.04.003

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786. doi: 10.3758/BF03196772

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*(1), 42-100. doi: 10.1016/j.cogpsych.2004.12.001

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671-684. doi: 10.1037/h0043943

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466. doi: 10.1016/S0022-5371(80)90312-6

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*(1), 13-21. doi: 10.1016/j.intell.2006.02.001

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods*, *17*(6), 652-655. doi: 10.3758/BF03200977

Draheim, C., Harrison, T. L., Embretson, S. E., & Engle, R. W. (2017). What item response theory can tell us about the complex span tasks. *Psychological Assessment, 30*(1), 115-129. doi:10.1037/pas0000444

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation, 44*, 145-200. doi: 10.1016/S0079-7421(03)44005-X

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331. doi: 10.1037/0096-3445.128.3.309

Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *Journal of Experimental Psychology: General*, *144*(5), 916-924. doi: 10.1037/xge0000101

Hooper, D., Coughlin, J., & Mullen, M. (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*, 53-60.

Jastrzębski, J., Ciechanowska, I., & Chuderski, A. (2018). The strong link between fluid intelligence and working memory cannot be explained away by strategy use. *Intelligence, 66*, 44-53. doi: 10.1016/j.intell.2017.11.002

Jarosz, A. F., & Wiley, J. (2012). Why does working memory capacity predict RAPM performance? A possible role of distraction. *Intelligence, 40*(5), 427-438.

doi: 10.1016/j.intell.2012.06.001

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130*(2), 169-183. doi: 10.1037/0096-3445.130.2.169

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Eds.), *Variation in working memory* (21-48). NY: Oxford University Press.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*(1), 66-71. doi: 10.1037/0033-2909.131.1.66

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

Loesche, P., Wiley, J., Hasselhorn, M. (2015). How knowing the rules affects solving the Raven Advanced Progressive Matrices Test. *Intelligence, 48*, 58-75. doi: 10.1016/j.intell.2014.10.004

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281. doi: 10.1038/36846

McNamara, D. S., & O'Reilly, T. (2009). Theories of comprehension skill: Knowledge and strategies versus capacity and suppression. In A. M. Columbus (Ed.), *Advances in psychology research* (113-136). Hauppauge, NY: Nova Science Publishers, Inc.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition, 29*(1), 10-17. doi: 10.3758/BF03195736

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review, 11*(2), 296-301. doi: 10.3758/BF03196573

Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*(4), 1343-1355. doi: 10.3758/s13428-014-0543-2

Raven, J. C. (1936). *Mental tests used in genetic studies: The performances of related individuals in tests mainly educative and mainly reproductive* (Thesis). University of London.

Raven, J. C. (1941). Standardization of progressive matrices, 1938. *Psychology and Psychotherapy: Theory, Research and Practice*, *19*(1), 137-150. doi: 10.1111/j.2044-8341.1941.tb00316.x

Raven, J. C. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41*(1), 1-48. doi: 10.1006/cogp.1999.0735

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales. Sections 3 & 4.* San Antonio, TX: The Psychological Corporation.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5/6*, 359–371. doi: 10.1111/j.1751-9004.2011.00355.x

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4-27. doi: 10.1037/0096-3445.125.1.4

Shipstead, Z., Harrison, T., & Engle, R. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception and Psychophysics, 77*(6), 1863-1880.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15*(2), 201-292.

Thorsen, C., Gustafsson, J., & Cliffordson, C. (2014). The influence of fluid and crystallized intelligence on the development of knowledge and skills. *British Journal of Educational Psychology, 84*(4), 556-570. doi: 10.1111/bjep.12041

Thurstone, L. L., & Thurstone, J. (1962). Test of primary mental abilities (Revised ed.). Chicago: Chicago Science Research Association.

Tormala, Z. L., Falces, C., Briñol, P., & Petty, R. E. (2007). Ease of retrieval effects in social judgment: The role of unrequested cognitions. *Journal of Personality and Social Psychology, 93*, 143–157. doi: 10.1037/0022-3514.93.2.143

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1*, 86 –98.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of memory and language, 28*(2), 127-154. doi: 10.1016/0749-596X(89)90040-5

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, *30*(2), 128-134. doi: 10.1037/h0076759

Unsworth, N. (2016). The many facets of individual differences in working memory capacity. In B. Ross (Ed.). *The Psychology of Learning and Motivation, 65,* 1-46.

Unsworth, N., & Brewer, G. A. (2009). Examining the relationships among item recognition, source recognition, and recall from an individual differences perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1578-1585. doi: 10.1037/a0017255

Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104-132. doi: 10.1037/0033-295X.114.1.104

Unsworth, N., & Engle, R. W. (2007b). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*(6), 1038-1066. doi: 10.1037/0033-2909.133.6.1038

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1-26. doi: 10.1016/j.cogpsych.2014.01.003

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods, 37*(3), 498-505. doi: 10.3758/BF03192720

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17*(6), 635-654. doi: 10.1080/09658210902998047

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence*, *38*(1), 111-122. doi: 10.1016/j.intell.2009.08.002

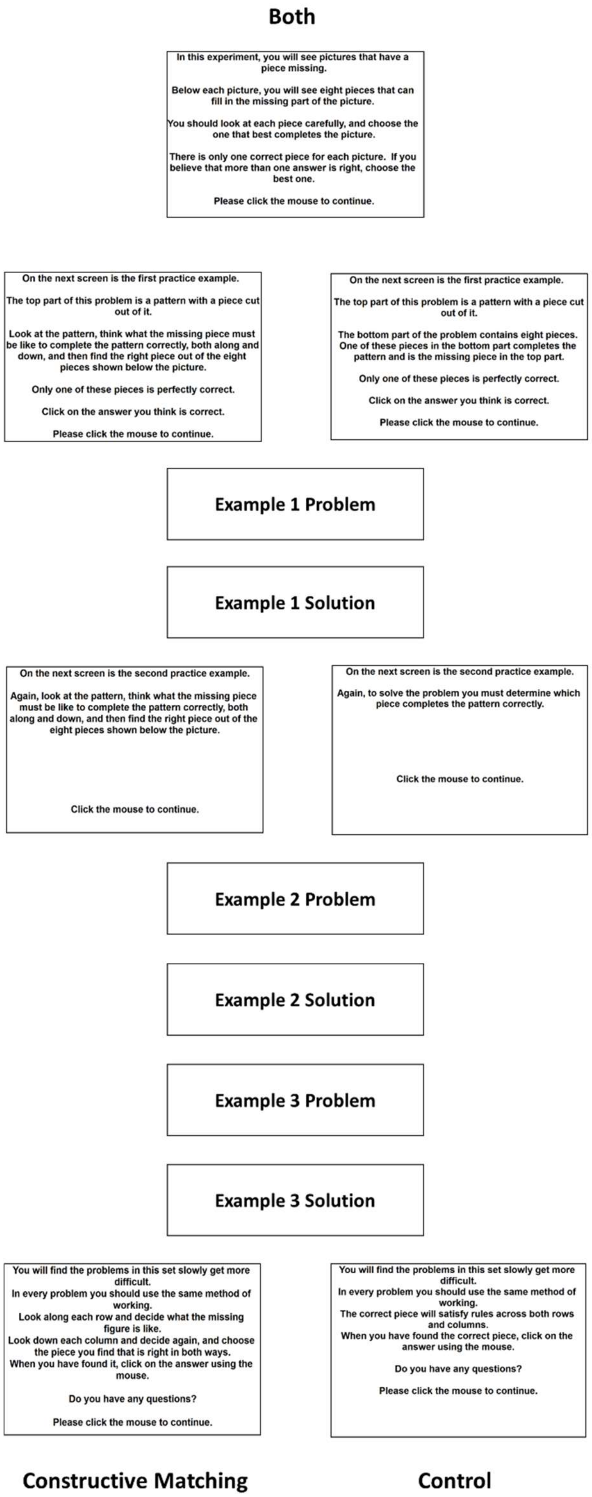
Unsworth, N., & Redick, T. S. (2017). Working memory and intelligence. *Reference module in neuroscience and biobehavioral psychology*. doi: 10.1016/B978-0-12-809324-5.21041-9.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2010). The contributions of primary and secondary memory to working memory capacity: An individual differences analysis of immediate free recall. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 36*(1), 240-247. doi: 10.1037/a0017739

Vigneau, F., Caissie, A. F., & Bors, D. A. (2006). Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence, 34*(3), 261-272. doi: 10.1016/j.intell.2005.11.003

APPENDIX A

MATERIALS USED IN EXPERIMENTS 1 – 3



In this experiment, you will see pictures that have a piece missing.

Below each picture, you will see eight pieces that can fill in the missing part of the picture.

You should look at each piece carefully, and choose the one that best completes the picture.

There is only one correct piece for each picture. If you believe that more than one answer is right, choose the best one.

Please click the mouse to continue.

On the next screen is the first practice example.

The top part of this problem is a pattern with a piece cut out of it.

Look at the pattern, think what the missing piece must be like to complete the pattern correctly, both along and down, and then find the right piece out of the eight pieces shown below the picture.

Only one of these pieces is perfectly correct.

Click on the answer you think is correct.

Please click the mouse to continue.

On the next screen is the first practice example.

The top part of this problem is a pattern with a piece cut out of it.

The bottom part of the problem contains eight pieces. One of these pieces in the bottom part completes the pattern and is the missing piece in the top part.

Only one of these pieces is perfectly correct.

Click on the answer you think is correct.

Please click the mouse to continue.

On the next screen is the second practice example.

Again, look at the pattern, think what the missing piece must be like to complete the pattern correctly, both along and down, and then find the right piece out of the eight pieces shown below the picture.

Click the mouse to continue.

On the next screen is the second practice example.

Again, to solve the problem you must determine which piece completes the pattern correctly.

Click the mouse to continue.

You will find the problems in this set slowly get more difficult.
In every problem you should use the same method of working.
Look along each row and decide what the missing figure is like.
Look down each column and decide again, and choose the piece you find that is right in both ways.
When you have found it, click on the answer using the mouse.

Do you have any questions?

Please click the mouse to continue.

You will find the problems in this set slowly get more difficult.
In every problem you should use the same method of working.
The correct piece will satisfy rules across both rows and columns.
When you have found the correct piece, click on the answer using the mouse.

Do you have any questions?

Please click the mouse to continue.

Figure A. Instructions for the RAPM task used in Experiments 1 and 2.

Table A

Questions and question type for the Experiment 3 post-experimental questionnaire

Question	Type	Task
I took the time to examine the drawing and to think about the answer before examining the response alternatives.	Constructive Matching	RAPM
After examining the drawing, I imagined the missing piece and then looked for it among the possible answers.	Constructive Matching	RAPM
After examining the drawing, I ruled out the response alternatives that did not match until only one remained.	Response Elimination	RAPM
I successfully examined each possible answer to decide whether it could be the missing piece.	Response Elimination	RAPM
I took the time to examine the numbers and to think about the answer before examining the response alternatives.	Constructive Matching	Number Series
After examining the numbers, I imagined a possible solution and then looked for it among the possible answers.	Constructive Matching	Number Series
After examining the numbers, I ruled out the response alternatives that did not match until only one remained.	Response Elimination	Number Series
I successively examined each possible answer to decide whether it could be the missing piece.	Response Elimination	Number Series
I took the time to examine one of the letters and to think about the rule before checking it with the other response alternatives.	Constructive Matching	Letter Sets
After examining the letters, I imagined a possible solution and then looked for it among the possible answers.	Constructive Matching	Letter Sets
After examining the letters, I ruled out the response alternatives that did not match until only one remained.	Response Elimination	Letter Sets
I successively examined each possible answer to decide whether it could be the inappropriate set.	Response Elimination	Letter Sets