The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Individual differences in working memory capacity and resistance to belief bias in syllogistic reasoning

Matthew K. Robison & Nash Unsworth

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2017
VOL. 70, NO. 8, 1471–1484
http://dx.doi.org/10.1080/17470218.2016.1188406

Routledge
Taylor & Francis Group

Individual differences in working memory capacity and resistance to belief bias in syllogistic reasoning

Matthew K. Robison and Nash Unsworth

Department of Psychology, University of Oregon, Eugene, OR, USA

ABSTRACT

In two experiments, we investigated the possibility that individual differences in working memory capacity (WMC) would provide resistance to belief bias in syllogistic reasoning. In Experiment 1 ($N = 157$), participants showed a belief bias effect in that they had longer response times and decreased accuracy on syllogisms with conflict between the validity and believability of the conclusion than on syllogisms with no such conflict. However, this effect did not differ as a function of individual differences in WMC. Experiment 2 ($N = 122$) replicated this effect with the addition of decontextualized (i.e., nonsense) syllogisms as a different means of measuring the magnitude of the belief bias effect. Although individual differences in WMC and fluid intelligence were related to better reasoning overall, the magnitude of the belief bias effect was not smaller for participants with greater WMC. The present study offers two novel findings: (a) The belief bias effect is independent of individual differences in WMC and fluid intelligence, and (b) resolving conflict in verbal reasoning is not a type of conflict resolution that correlates with individual differences in WMC, further establishing boundary conditions for the role of WMC in human cognitive processes.

ARTICLE HISTORY
Received 29 October 2015
Accepted 5 May 2016
First Published Online 8 June 2016

KEYWORDS
Working memory; reasoning; biases

The ability to reason logically is an important cognitive skill that we call upon to make decisions about information with which we are presented. Syllogisms are often used as a measure of reasoning abilities (e.g., Sá, West, & Stanovich, 1999; Stanovich & West, 1997). When presented with a syllogism, participants are asked to determine the validity of the conclusion based on a set of premises, operating under the assumption that the premises are true. Sometimes, the conclusion is consistent with an individual's beliefs, or perhaps may seem factually true. Other times, the conclusion may actually conflict with an individual's prior beliefs. In such cases, people must reason independently of prior belief and instead focus solely on the logic presented by the argument. At other times a conclusion may be in line with an individual's prior beliefs, but the conclusion is not necessitated by the premises. In such cases people must also reason independently of prior belief. The tendency to rely on prior beliefs rather than strict adherence to logical principles is known as belief bias, and susceptibility to belief bias can significantly alter one's ability to reason.

Belief bias can lead individuals to incorrectly endorse a conclusion or refute strong evidence that runs against an individual's beliefs and as a result can hinder rational thought. As an example, a doctor may have a particular idea about the proper diagnosis for a patient, but if this belief leads the doctor to ignore the deductive process in properly ruling out alternative possibilities, a misdiagnosis can occur. In another instance, a juror may develop a belief about a defendant, and instead of following the logic of the arguments by the defence and the prosecution, the juror bases their conclusion on that belief. As these examples indicate, belief bias can have serious real-world implications, and therefore it is an important area of inquiry.

From the dual-process framework of thinking (Kahneman, 2011; Stanovich, 1999), people reason

CONTACT Matthew K. Robison mkr@uoregon.edu Department of Psychology, 1227 University of Oregon, Eugene, OR 97403, USA

using two broad systems of thought. System 1 is fast, intuitive, and relatively automatic, and relies on heuristics. System 2 is slower, more analytical, and controlled, and tries to resist any heuristics or biases. People can apply both of these systems of thinking, but one will be more advantageous than the other depending on the context. If the conclusion is valid based on the argument, and it aligns with an individual's beliefs, System 1 will work perfectly fine in arriving at the accurate answer. Similarly if the conclusion is invalid, and the individual does not agree with the conclusion, System 1 will also be more effective, as it will arrive at the correct answer quickly. However, in situations in which there is conflict between the believability and validity of the conclusion, people must override System 1 processes in favour of System 2 processes to arrive at the appropriate answer. The present study is interested in individual differences in susceptibility to belief bias (i.e., the tendency to accept a believable, invalid conclusion and to reject an unbelievable, valid conclusion) and how those individual differences relate to working memory capacity (WMC) and fluid intelligence.

There are clear individual differences in verbal reasoning abilities, and these abilities correlate moderately with other measures of reasoning such as statistical reasoning and SAT scores, as well as non-cognitive personality measures like thinking disposition questionnaires (Stanovich & West, 1997, 1998, 2008). Stanovich and West (2008) argue that sources of individual differences in reasoning can arrive at a variety of stages in the reasoning process (see also, Kahneman, 2000). A "mindware gap" is an error of comprehension. In the case of syllogistic reasoning, some people may simply misunderstand the task and instead report that a conclusion is valid or invalid based solely on its apparent truth or falsehood. If all people have the relevant mindware (e.g., by giving people explicit task instructions), the next source of individual differences is the ability to detect that a System 1 response must be overridden. If people do not detect that an override is necessary, they will emit a heuristic response. In the next stage, people must decouple and decontextualize information in the service of overriding the System 1 response. Finally, people must have the capacity to sustain the decoupling process and sustain that override. In all of these stages, individual differences can arise. Based on task variation, instructions, and cognitive abilities of participants, among other variables, individual differences can arise from preferential use of System 1, or alternatively failures of System 2, in one or any combination of these stages.

Belief bias can also arise at various stages in the reasoning process. People can simply rely on prior beliefs to inform their decisions about the validity of a conclusion, they can fail to detect that their beliefs run counter to logical reasoning in a particular case, and they can recognize the need to override their prior beliefs but be unable to sustain such an override (e.g., due to low cognitive capacity). In the present study, we wanted to determine whether measures of WMC would predict individual differences in susceptibility/resistance to belief bias.

WMC is the ability to maintain and manipulate information, often in the presence of distraction, and is a core cognitive construct. WMC correlates moderately with measures of reasoning like the Raven matrices (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Raven, Raven, & Court, 1998; Unsworth, 2014), analogies (Kane et al., 2004; Unsworth, 2014), and number series (Unsworth, 2014; Unsworth, Fukuda, Awh, & Vogel, 2014) and has been shown to correlate with syllogistic reasoning as well (Copeland & Radvansky, 2004; De Neys, Schaeken, & D'Ydewalle, 2005; Kane et al., 2004; Markovits, Doyon, & Simoneau, 2002). At the latent level, WMC and reasoning correlate around .60–.70 (Unsworth, 2014; Unsworth, Redick, Heitz, Broadway, & Engle, 2009). The reason behind the shared variance between WMC and syllogistic reasoning is still not entirely clear. De Neys et al. (2005) argue that individuals with greater WMC are able to use this resource to develop relevant counter-examples. Copeland and Radvansky (2004) argue that WMC gives people the ability to build mental models of the syllogism (Johnson-Laird, 2010). They found that when people had to develop more than one mental model, the correlation between WMC and syllogistic reasoning increased. Markovits et al. (2002) also argue that WMC is related to reasoning because it allows individuals to develop models of the arguments and thus manipulate them to come to a correct answer. However, it is still not entirely clear whether or not WMC also correlates with resistance to belief bias.

Some prior work has observed a relationship between cognitive abilities and belief bias. Stanovich and West (1998) found a correlation between cognitive ability and belief bias in syllogistic reasoning, but their measures of cognitive ability did not specifically measure WMC. Their cognitive ability composite was a sum of the standardized scores on Raven matrices, Scholastic Assessment Test (SAT) scores, and the Nelson–Denny reading comprehension test. Although these measures do correlate moderately with WMC, they do not specifically tap the latent construct of WMC, which could be a source of the computational ability to detect the need to override System 1 and effectively employ System 2 in the service of reasoning despite conflict between beliefs and logic. Another aspect of Stanovich and West (1998) that makes the belief bias–cognitive ability relationship unclear is their exclusive use of syllogisms in which the validity and believability of the conclusion conflicted. Therefore, the observed correlation could be due to baseline reasoning ability rather than a specific resistance to belief bias. However Sá et al. (1999) observed a correlation between cognitive ability and belief bias as measured by a difference in accuracy between consistent (match between validity and believability) and inconsistent syllogisms.

In a separate study, Quayle and Ball (2000) measured WMC with a spatial span task and an articulatory span task, and measured belief bias with syllogisms. They found an interaction between belief bias and span, but only for the spatial span. Additionally, Quayle and Ball (2000) had only 32 participants in the experiment that found such an effect and utilized a median split to separate individuals into high- and low-WMC groups. De Neys et al. (2005) measured conditional reasoning with high- and low-WMC participants. To measure the effects of belief bias, De Neys et al. included conclusions with few and many possible alternatives or disablers. De Neys et al. argue that high-WMC participants are better at conditional reasoning because they are better able to search for alternatives and disablers. When their working memory is disrupted with an interfering dual task, high-WMC individuals are presumably hampered in their ability to generate alternatives and disablers, which led to declining accuracy for both valid and invalid conclusions. In a subsequent study, De Neys (2006) measured a broader range of participants and gave participants conflict and non-conflict syllogisms under conditions of high cognitive load, low cognitive load, and no load. De Neys found an interaction between span and syllogism type, such that WMC was only related to performance on the conflict syllogisms. However, because of the simultaneous effect of load, it is still not entirely clear whether WMC offered a specific resistance to belief bias. Specifically, in the no-load condition, it is not apparent that high-WMC participants differed from low- or mid-WMC participants more so for conflict than for non-conflict syllogisms. So when taken together these results do not provide a definitive answer or explanation for the role of WMC in resisting belief bias. However, there are additional findings that suggest that belief bias may be driving the WMC–syllogistic reasoning relationship.

Under the two-factor theory of cognitive control (Engle & Kane, 2004), the two factors that lead to individual differences in WMC, and thus their resulting relationships with other cognitive constructs like attention control, are goal-maintenance and conflict resolution. The ability to maintain and execute the task goal, often in the presence of interference, is one source of individual differences. In the present case, the task goal is to determine the validity of the conclusion, but in the case of conflict between validity and believability, there is interference that the participant must overcome. Emitting the heuristic response would be a failure of goal maintenance. Alternatively, WMC may offer resistance to belief bias as a result of individual differences in conflict resolution. In the present study, participants faced conflict on the two types of conflict syllogisms and had to resolve that conflict in favour of rationality. Therefore, the two-factor theory of WMC offers another reason for its potential to explain individual differences in belief bias. However, it remains an open question whether WMC can offer resistance to cognitive biases, and we sought to investigate the relationship between WMC and belief bias from an individual differences perspective.

The present study

The goal of the present study is to examine the WMC–reasoning relationship by investigating whether individual differences in WMC predict individual differences in susceptibility to belief bias. If WMC is the source of the computational limitation in the ability to override beliefs and reason independently of those beliefs, then individuals with greater WMC should show a reduced belief bias effect. If this is the case, resistance to belief bias could be one of the sources of the WMC–reasoning relationship. If, however, the source of belief lies within a different cognitive ability, or perhaps rather in a dispositional difference, there should be no relationship between belief bias effects and WMC. In that case, differential susceptibility to belief bias would not be one of the reasons behind the WMC–reasoning relationship.

The belief bias effect can be measured in two ways: accuracy and response time (RT). A difference in accuracy for syllogisms with consistency between the validity and believability of the conclusion (valid/believable, invalid/unbelievable) and for syllogisms with conflict between validity and believability (valid/unbelievable and invalid/believable) can be used as a measure of belief bias. If there is a relationship between the size of this effect and WMC, then we can conclude that working memory resources are the source of individual difference in overriding belief bias. However, it is also possible that this effect does not relate to WMC, and rather the difference arises in RT. Because System 2 processes are slow, it is possible that individuals with greater WMC are not more likely to employ this system to override belief bias, but rather they are more efficient. In this case, individuals with greater WMC will show a reduced belief bias effect in their RT, but no such effect in accuracy. Of course, both can arise. An additional possibility is that low-capacity individuals are simply more likely to emit the heuristic response and will thus show a reduced RT difference between conflict syllogisms and no-conflict syllogisms. A final possibility is that there is no relationship between WMC and belief bias. In other words, the ability to override belief bias is an independent cognitive ability and not a manifestation of WMC. In this case, there should be a main effect of WMC in that individuals with higher WMC show better reasoning, but this would not interact with the various types of syllogisms (conflict vs. no-conflict).

Experiment 1

The first experiment investigated individual differences in WMC, fluid intelligence (gF), and syllogistic reasoning. The primary goal of the experiment was the measure reasoning abilities and susceptibility to belief bias, as well as to see how individual differences in WMC potentially offered resistance to belief bias. To do this, we manipulated both believability and validity of conclusions in an entirely within-subjects design.

Method

Participants

A sample of 157 participants (110 females) from the University of Oregon undergraduate subject pool participated in partial fulfilment of a course requirement. All participants were between the ages of 18 and 41 years ($M = 19.57$, $SD = 2.70$). All participants gave informed consent and were debriefed following completion of the study. Minimum target sample size was 120 with the end of the academic term as our stopping rule for data collection.

Procedure

After giving informed consent, participants completed three tasks measuring working memory capacity (operation span, symmetry span, and reading span), two tasks measuring fluid intelligence (letter sets and number series), and a syllogistic reasoning task. Participants also completed measures of visual working memory and attention control, but because they are not the focus of the current investigation, they are not reported here. Experimental sessions lasted two hours.

Tasks

Working memory capacity

Operation span. The span tasks were used to measure working memory capacity because they require participants to both process and store information in working memory. In this task, participants solved a series of maths operations while trying to remember a set of unrelated letters. Participants were required to solve a maths operation, and after solving the operation, they were presented with a letter for 1 s. Immediately after the letter was presented the next operation was presented. At recall participants were asked to recall letters from the current set in the correct order by clicking on the appropriate letters. For all of the span measures, items were scored correct if the item was recalled correctly from the current list in the correct serial position. Participants were given practice on the operations and letter recall tasks only, as well as two practice lists of the complex, combined task. List length varied randomly from three to seven items, and there were two lists of each length for a total possible score of 50. The score was total number of correctly recalled items in the correct serial position.

Symmetry span. Participants recalled sequences of red squares within a matrix while performing a symmetry-judgment task. In the symmetry-judgment task, participants were shown an 8×8 matrix with some squares filled in black. Participants decided whether the design was symmetrical about its vertical axis.

The pattern was symmetrical half of the time. Immediately after determining whether the pattern was symmetrical, participants were presented with a 4 × 4 matrix with one of the cells filled in red for 650 ms. At recall, participants recalled the sequence of red-square locations by clicking on the cells of an empty matrix. Participants were given practice on the symmetry-judgment and square recall task as well as two practice lists of the combined task. List length varied randomly from two to five items, and there were two lists of each length for a total possible score of 28. We used the same scoring procedure as the one that we used in the operation span task.

Reading span. While trying to remember an unrelated set of letters, participants were required to read a sentence and indicated whether or not it made sense. Half of the sentences made sense, while the other half did not. Nonsense sentences were created by changing one word in an otherwise normal sentence. After participants gave their response, they were presented with a letter for 1 s. At recall, participants were asked to recall letters from the current set in the correct order by clicking on the appropriate letters. Participants were given practice on the sentence judgment task and the letter recall task, as well as two practice lists of the combined task. List length varied randomly from three to seven items, and there were two lists of each length for a total possible score of 50. We used the same scoring procedure as the one that we used in the operation span and symmetry span tasks.

Fluid intelligence.

Number series. In this task, participants saw a series of numbers and were required to determine what the next number in the series should be (Thurstone, 1962). That is, the series follows some unstated rule, which participants are required to figure out in order to determine what the next number in the series should be. The ability to detect these patterns is a commonly noted element of fluid intelligence. Participants selected their answer from five possible numbers that were presented. Following five practice problems, participants had 3.5 min to complete 15 test times. A participant's score was the total number of items solved correctly.

Raven Advanced Progressive Matrices. The Raven is a measure of abstract reasoning and is commonly used in intelligence batteries (Raven et al., 1998). The test consists of 36 items presented in ascending order of difficulty (i.e., easiest to hardest). Each item consists of a display of 3 × 3 matrices of geometric patterns with the bottom right pattern missing. The task for the participant is to select among eight alternatives the one that correctly completes the overall series of patterns. Participants received two practice items and were then given 10 min to complete the 18 odd-numbered items. A participant's score was the total number of correct solutions.

Syllogistic reasoning. Participants were instructed to determine the validity of a conclusion based on a set of two premises. Specifically, the instructions said "You should only state that the argument is valid if the conclusion follows necessarily from the premises. You should NOT determine whether or not you think the conclusion is true. For all arguments, you should assume both the premises are true. If the conclusion necessarily follows from the truth of the premises, you should indicate that the argument is valid. If not, indicate the argument is invalid." These instructions attempted to control for varying task construals among participants. Participants were then shown an example of a valid argument and an example of an invalid argument. They were to press a key marked "I" for invalid or a key marked "V" for valid. The primary dependent variables of interest were accuracy and reaction time to the syllogisms. All valid syllogisms were presented in the following form: All X are Y; all Z are X; therefore, all Z are Y. All invalid syllogisms were presented in the following form: All X are Y; all X are Z; therefore, all Y are Z. On half of the valid syllogisms and half of the invalid syllogisms, the conclusions were believable, and on the other half the conclusion was not believable.[1] Participants were given four syllogisms of each type: valid/believable conclusion; valid/unbelievable conclusion; invalid/believable conclusion; and invalid/unbelievable conclusion. Syllogisms were presented in a different random order for each participant. An example of each type is shown in Table 1.

Results

We first standardized the scores on the operation span, symmetry span, and reading span and averaged the *Z*-scores to give each participant a single WMC score. This score was used in all subsequent analyses involving WMC. Descriptive statistics for the complex span tasks, Raven, and letter sets are shown in Table 2. For correlations between WMC, measures of fluid intelligence, and syllogistic reasoning, see Table 4.

Table 1. Examples of syllogisms of each type.

Valid	Believable	Syllogism
Yes	Yes	All candy is made out of sugar. All lollipops are candy. Therefore, all lollipops are made out of sugar.
Yes	No	All doctors have medical degrees. All college professors are doctors. Therefore, all college professors have medical degrees.
No	Yes	All sea creatures are animals that are able to swim. All sea creatures are animals that spend the majority of their lives in the water. Therefore, all animals that are able to swim spend the majority of their lives in the water.
No	No	All cubes are objects with six sides. All cubes are objects with sides of equal area. Therefore, all objects with sides of equal area have six sides.

Table 2. Descriptive statistics for WMC and gF measures in Experiment 1.

Measure	Mean	*SD*	Range	Skew	Kurtosis
Operation span	36.85	9.16	2–50	−0.89	0.75
Symmetry span	19.57	5.28	5–28	−0.54	−0.21
Reading span	36.12	9.50	3–50	−0.93	1.22
RAPM	8.95	2.88	2–18	−0.01	0.06
Letter sets	10.11	3.09	4–18	0.22	−0.61

Note: $N = 157$. WMC = working memory capacity; gF = fluid intelligence; RAPM = Raven Advanced Progressive Matrices.

Accuracy

We submitted proportion correct for each of the four types of syllogisms to a 2 × 2 repeated measures analysis of covariance (ANCOVA) with validity (valid/invalid) and believability (believable/unbelievable) as within-subjects factors and WMC as a covariate. The ANCOVA revealed a main effect of validity, $F(1, 155) = 18.40$, $p < .001$, $\eta^2_p = .10$, indicating that participants were more accurate on valid syllogisms, but no significant main effect of believability ($p = .19$). However there was a significant interaction between validity and believability, $F(1, 155) = 210.84$, $p < .001$, $\eta^2_p = .57$, such that participants scored lowest when there was a conflict between the validity and believability of the conclusion. Mean accuracy for each of the four types of syllogisms are listed in Table 3. WMC showed a main effect on accuracy, $F(1, 155) = 8.85$, $p < .01$, $\eta^2_p = .05$, which indicates that overall individuals with greater WMC were more accurate on the syllogisms. Collapsed across syllogism types, WMC and accuracy significantly correlated ($r = .23$, $p < .01$). However, neither the main effect of validity ($p = .23$) nor the interaction between validity and believability ($p = .25$) showed a significant interaction with WMC.[2]

We categorized individuals as high WMC and low WMC based on a quartile split. Participants in the top quartile for WMC were labelled as "high WMC", and participants in the lowest quartile were labelled as "low WMC". For illustrative purposes, we show only high- and low-WMC groups. The pattern of results in Figure 1 indicates that although greater WMC is related to better reasoning, resistance to belief bias does not moderate this effect.

As a final test, we computed a measure of belief bias for valid and invalid syllogisms separately. To do so, we computed the difference in accuracy on valid/believable syllogisms and valid/unbelievable syllogisms, as well as the difference in accuracy for invalid/believable syllogisms and invalid/unbelievable syllogisms. The bias effect on accuracy did not significantly correlate with WMC for either valid syllogisms ($r = -.07$, $p = .34$) or invalid syllogisms ($r = -.06$, $p = .43$). But, as noted earlier, it is possible that the effect of WMC is expressed in response time, rather than accuracy.

Response time

We submitted response times (RTs)[3] to a 2 × 2 repeated measures ANCOVA with validity and believability again as the within-subjects factors and WMC as a covariate. Because the syllogism types varied in their average length, we used RT per syllable as the dependent variable in the analyses. The ANCOVA revealed a main effect of validity, $F(1, 155) = 30.65$, $p < .001$, $\eta^2_p = .16$, indicating that participants were faster to respond to valid syllogisms than to invalid syllogisms, and a main effect of believability, $F(1, 155) = 409.75$, $p < .001$, $\eta^2_p = .72$, indicating that participants were faster to respond to syllogisms with believable conclusions than to those with unbelievable conclusions. Importantly, these two factors interacted, $F(1, 155) = 51.01$, $p < .001$, $\eta^2_p = .25$, indicating that participants were slower when the validity and the believability of the conclusion conflicted. In this case, WMC did not show a main effect of RT, indicating that individuals with greater WMC did not respond faster to the syllogisms, overall. Just as with accuracy, none of the observed effects interacted with WMC (all *p*s > .20), indicating that this pattern of results did not differ across individuals with varying WMC estimates[4] (Figure 2). Again as a final test, we computed the difference in RT for valid syllogisms (valid/unbelievable – valid/believable) and invalid syllogisms (invalid/believable – invalid/unbelievable). Neither the effect for valid syllogisms ($r = .06$, $p = .43$) nor the effect for

Table 3. Mean accuracy and response time for each type of syllogism in Experiment 1.

Validity	Believability	Accuracy	RT/syllable
Valid	Believable	.87 (.17)	358.95 (115.18)
Valid	Unbelievable	.62 (.34)	546.51 (220.43)
Invalid	Believable	.50 (.27)	344.29 (122.69)
Invalid	Unbelievable	.89 (.25)	697.99 (311.76)

Note: RT = response time. Numbers in parentheses are standard deviations. Accuracy is proportion correct, and response times are in seconds.

invalid syllogisms ($r = .04$, $p = .62$) correlated with WMC. Together, the results indicate that WMC does not offer resistance to the belief bias effect.[5]

Participants had the most difficulty determining the validity of a syllogism when the validity and the believability of the conclusion did not match. However, it does not appear that participants were simply using System 1 processes in responding to these syllogisms as they took the longest amount of time on this type. Rather, it seems as though they were using their System 2 processing, but it still did not always bring them to the correct conclusion. However, it could be the case that some participants were indeed using effective System 2 processes and coming to the correct conclusion while others were susceptible to belief bias and reverted to System 1 processes and gave the pre-potent response of "valid" to an invalid syllogism with a believable conclusion.

To investigate this possibility, we examined correlations between response times and accuracy for the four different types of syllogisms. If participants are able to use System 1 processes when the logic of the problem and believability of the conclusion match (valid/believable, invalid/unbelievable), then there should not be a correlation between response time and accuracy on these syllogisms as both quick System 1 processes and more time-consuming System 2 processes can both arrive at the correct response. Alternatively, there should be a correlation between response time and accuracy on "conflict" syllogisms (valid/unbelievable, invalid/believable). This was indeed the case. Response time and accuracy did not correlate for match syllogisms ($r = .03$, $p = .71$), but did correlate for conflict syllogisms ($r = .32$, $p < .01$). This effect seemed to be driven mostly by performance on the valid/unbelievable syllogisms, as this was the only type of syllogism that showed a significant difference in response time between accurate and inaccurate responses, $t(616.25) = 4.05$, $p < .001$ (all other ps $> .06$).[6] This indicates that more controlled, slower System 2 processes are needed to resolve conflict between prior beliefs and the logical solution, particularly in the case in which the conclusion is valid, yet unbelievable. But when the logic of the solution and prior beliefs are in alignment, both System 1 and System 2 processes will arrive at the correct solution, and the lack of a correlation between response time and accuracy for these types of syllogisms indicates that people are differentially using these systems to solve these problems.

Discussion

The goal of Experiment 1 was to test whether resistance to belief bias was one potential explanation for individual differences in reasoning and potentially a factor in

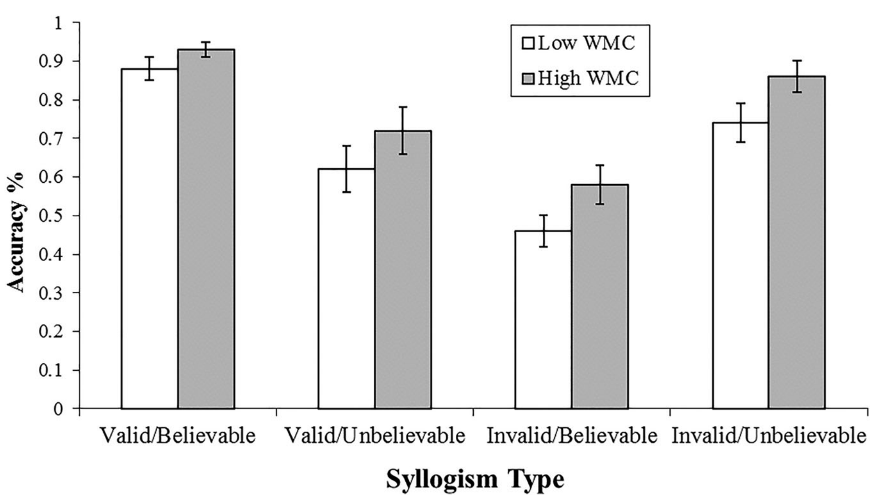

Figure 1. Accuracy for high- and low-WMC (working memory capacity) participants for each syllogism type in Experiment 1.

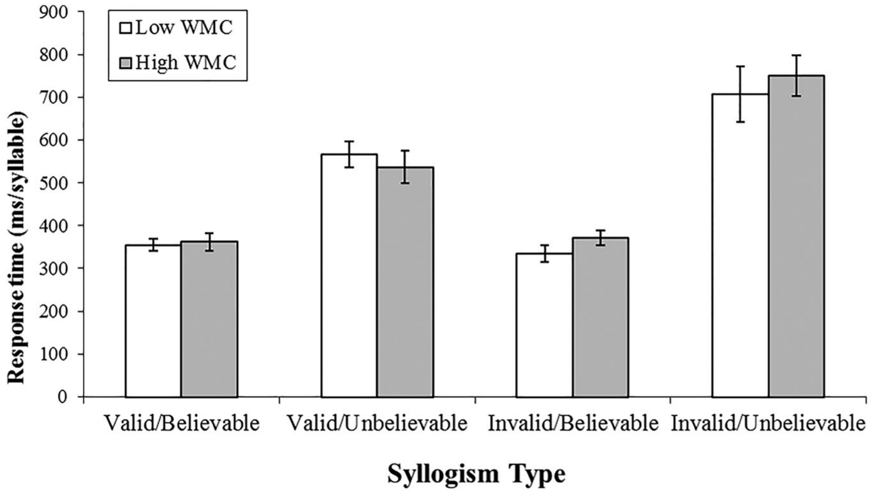

Figure 2. Response times for high- and low-WMC (working memory capacity) participants for each syllogism type in Experiment 1.

the WMC–reasoning relationship. Despite evidence that participants had to override the pre-potent response in syllogisms in which there is a conflict between the validity and believability of the conclusion, as indexed by decreases in accuracy and longer response times, these effects did not interact with WMC. In other words, individuals with greater WMC, and thus a greater capacity to perform mental operations, did not show a decreased belief bias effect. Therefore, although WMC was related to better reasoning ability overall, we cannot conclude that WMC is the driving factor behind resistance to belief bias.

Experiment 2

The goal of Experiment 2 was to replicate the findings of Experiment 1 with one additional analysis. We added syllogisms with nonsense words to measure baseline reasoning abilities free of any context. We wanted to ensure that the resistance to belief bias, although not affected by WMC as Experiment 1 showed, was not simply due to baseline verbal reasoning abilities. All analyses are nearly identical to those in Experiment 1. As a preview of the results, Experiment 2 largely replicated Experiment 1.

Method

Participants

A total of 122 participants (73 females) from the undergraduate subject pool at the University of Oregon participated in partial fulfilment of a course requirement. All participants were between the ages of 18 and 48 years ($M = 20.23$, $SD = 3.58$). All participants gave informed consent and were debriefed following completion of the study. Minimum target sample size was 120 with the end of the academic term as our stopping rule for data collection.

Table 4. Correlations among syllogistic reasoning, WMC, and gF in Experiment 1.

Measure	1	2	3
1. WMC	—		
2. Syllogisms	.23**	—	
3. RAPM	.40**	.28**	—
4. Letter sets	.40**	.24**	.24**

Note: $N = 157$. WMC = working memory capacity composite score; gF = fluid intelligence; syllogisms = mean accuracy on the syllogistic reasoning task; RAPM = Raven Advanced Progressive Matrices.
**$p < .01$.

Procedure

After giving informed consent, participants completed the same measures as those in Experiment 1 with one slight change: the addition of nonsense syllogisms. Participants also completed one measure of long-term memory and several measures of visual working memory, but because those measures were irrelevant to the current study, they are not reported here. Experimental sessions lasted two hours.

Tasks

Working memory capacity.
Operation span. See Experiment 1

Symmetry span. See Experiment 1.
Reading span. See Experiment 1.
Fluid intelligence.
Raven Advanced Progressive Matrices. See Experiment 1.
Letter sets. See Experiment 1.
Syllogistic reasoning. We added two valid and two invalid syllogisms with nonsense words (e.g., all weebles are greebles). The valid and invalid syllogisms followed the same form as the other syllogisms. Items were presented in a random order.

Table 5. Descriptive statistics for WMC and gF measures in Experiment 2.

Measure	Mean	SD	Range	Skew	Kurtosis
Operation span	38.02	8.02	6–50	−1.12	1.74
Symmetry span	19.63	4.73	6–28	−0.37	−0.22
Reading span	38.08	7.21	14–50	−0.72	0.40
RAPM	8.98	2.90	2–18	−0.01	0.004
Letter sets	10.28	3.03	4–17	0.04	−0.53

Note: $N = 122$. WMC = working memory capacity; gF = fluid intelligence; RAPM = Raven Advanced Progressive Matrices; *SD* = standard deviation.

Table 6. Mean accuracy and response time for each type of syllogism in Experiment 2.

Validity	Believability	Accuracy	RT/syllable
Valid	Believable	.81 (.18)	395 (148)
Valid	Unbelievable	.71 (.34)	561 (244)
Invalid	Believable	.53 (.28)	375 (193)
Invalid	Unbelievable	.82 (.28)	349 (150)
Valid	Decontextualized	.77 (.24)	903 (430)
Invalid	Decontextualized	.54 (.37)	870 (363)

Note: RT/syllable = response time per syllable in milliseconds. Numbers in parentheses are standard deviations. Accuracy is proportion correct, and response times are in milliseconds per syllable.

Results

Descriptive statistics for the WMC and gF measures are shown in Table 5. Just as in Experiment 1, we created a composite WMC score by standardizing scores on the complex span tasks and averaging these scores to give each participant a single WMC score.

Accuracy

Mean accuracies and standard deviations for each of the four syllogism types are shown in Table 6. We submitted the accuracy data to a repeated measures ANCOVA with within-subjects factors of validity (valid vs. invalid) and believability (believable vs. unbelievable), and we entered WMC into the model as a covariate. The ANCOVA revealed a main effect of WMC, $F(1, 119) = 5.60$, $p < .05$, $\eta_p^2 = .04$, suggesting that overall, participants with better WMC were more accurate on the syllogistic reasoning task. Collapsed across all syllogism types, accuracy correlated with WMC ($r = .31$, $p < .001$). The ANCOVA also revealed a main effect of validity, $F(1, 119) = 11.85$, $p < .05$, $\eta_p^2 = .09$, which indicates that participants were more accurate on valid syllogisms than on invalid syllogisms; a main effect of believability, $F(1, 119) = 21.53$, $p < .01$, $\eta_p^2 = .15$, which indicates that participants were more accurate on believable syllogisms than on unbelievable syllogisms; and an interaction between validity and believability, $F(1, 119) = 96.21$, $p < .01$, $\eta_p^2 = .44$, which indicates that participants were least accurate when the validity and believability of the conclusion did not match one another.

The validity effect was replicated with the decontextualized (i.e., nonsense) syllogisms. Participants were more accurate on valid syllogisms ($M = .77$, $SD = .24$) than on invalid syllogisms ($M = .54$, $SD = .37$), $t(121) = 5.68$, $p < .001$. However, the main effect of WMC did not reach significance for these types of syllogisms ($p = .50$).

Importantly, neither the main effects for validity and believability nor the interaction between validity and believability interacted with WMC ($p = .87$), suggesting that individuals with high WMC are not less susceptible to belief bias in reasoning. This was also the case when examining the validity effect for the decontextualized syllogisms ($p = .34$). Again, for illustrative purposes we plotted accuracy for each type of syllogism for high- and low-WMC participants (Figure 3). Using the same belief bias metrics as those in Experiment 1, neither the effect on accuracy for valid syllogisms ($r = .04$, $p = .61$) nor the effect for invalid syllogisms ($r = −.13$, $p = .14$) correlated with WMC.

Response times

We also examined response times (RTs) as a function of validity, believability, and WMC. We ran a similar repeated measures ANCOVA with validity and believability as within-subjects factors and WMC as a covariate. Mean RTs and standard deviations are listed in Table 6. There was no main effect for WMC ($p = .68$), suggesting that individuals with greater WMC were not faster to respond to the syllogisms overall. However, we again observed a main effect of validity, $F(1, 120) = 82.34$, $p < .001$, $\eta_p^2 = .40$, which

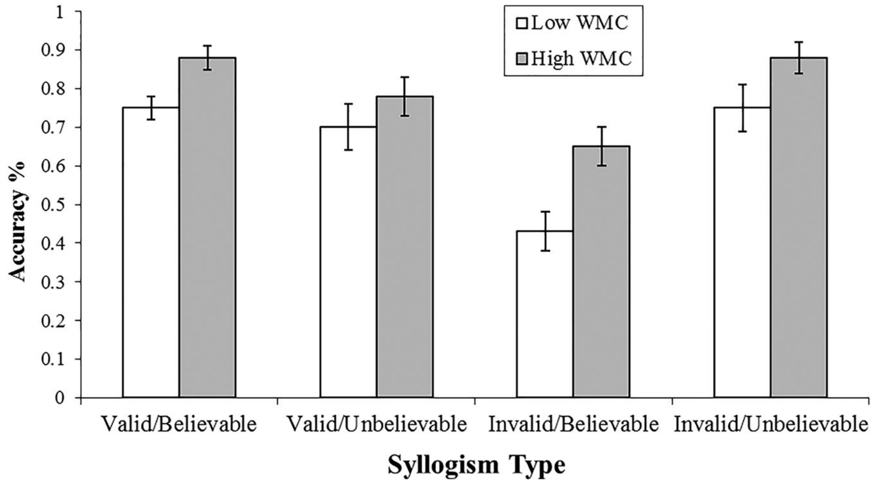

Figure 3. Accuracy for high- and low-WMC (working memory capacity) participants for each syllogism type in Experiment 2.

indicated that participants were faster to respond to valid syllogisms than to invalid syllogisms, and a main effect of believability, $F(1, 120) = 34.92$, $p < .001$, $\eta_p^2 = .22$. But validity and believability interacted, suggesting that participants were slowest to respond when validity and believability conflicted, $F(1, 120) = 60.13$, $p < .001$, $\eta_p^2 = .33$. Importantly, neither of the main effects nor the interaction interacted with WMC (all *p*s > .60). These results are depicted in Figure 4. Again using RT differences for valid and invalid syllogisms separately, neither the belief bias effect for valid syllogisms ($r = -.04$, $p = .65$) nor the effect for invalid syllogisms ($r = -.04$, $p = .65$) correlated with WMC. This suggests that although participants took longer to respond to syllogisms when they had to override the pre-potent response and deal with interference between the validity and believability of the conclusion, this effect was not greater for individuals with lower WMC.

We again analysed the difference in RT for accurate and inaccurate trials. As in Experiment 1, the effect of belief bias on RT was strongest for valid/believable conclusions, as RTs for accurate responses to this syllogism type were significantly lower than those for inaccurate responses, $t(475.64) = 3.98$, $p < .001$, which suggests that System 2 processes were more effective in this context. The effect was marginally significant for valid/believable trials, $t(465.70) = 2.00$, $p = .05$, but in the opposite direction. Participants were more accurate when they were faster, so System 1 processes were actually more effective in this context.

Measuring belief bias with decontextualized syllogisms

As an additional test of the relationship between WMC and susceptibility to belief bias, we calculated the difference in accuracy between decontextualized syllogisms (both valid and invalid) and verbal syllogisms with a conflict between the validity and believability of the conclusion. To do this, we subtracted each participant's proportion correct on valid-decontextualized syllogisms from their proportion correct on valid syllogisms with unbelievable conclusions. We also calculated the accuracy difference between invalid-decontextualized syllogisms and invalid syllogisms with believable conclusions. Presumably this difference measures the effect of having to decontextualize information from its real-world meaning. Neither of these measures of belief bias correlated with either WMC or gF (all *r*s < .13 in magnitude, all *p*s > .14). See Table 7.

We also examined the belief bias effect by subtracting RTs on decontextualized syllogisms from syllogisms with a conflict between the validity and believability of the conclusion. Again neither of these measures (i.e., the belief bias effects for valid and invalid syllogisms) correlated with either WMC or gF (all *p*s > .12). Therefore, we can conclude that although individuals with greater WMC and better fluid reasoning are better at reasoning through verbal syllogisms, there is no evidence that either of these abilities offers an advantage in resisting the effects of belief bias.

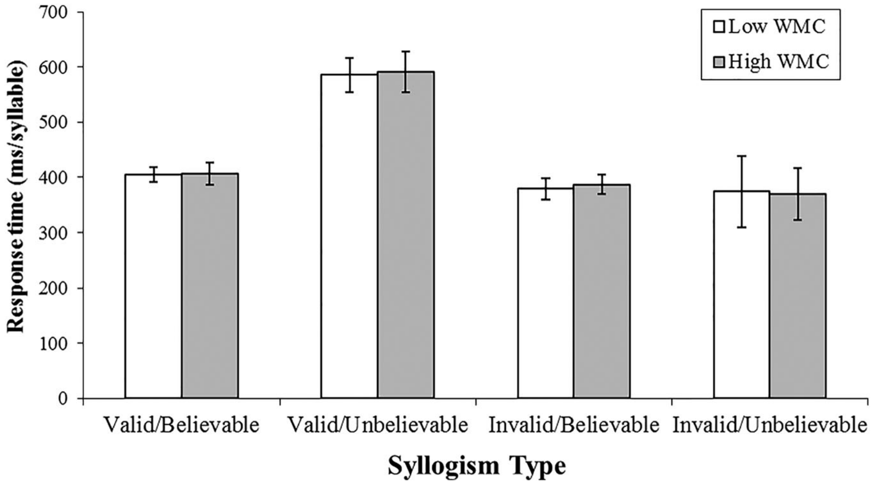

Figure 4. Response times for high- and low-WMC (working memory capacity) participants for each syllogism type in Experiment 2.

Table 7. Correlations among syllogistic reasoning, WMC, and gF in Experiment 2.

Measure	1	2	3
1. WMC	—		
2. Syllogisms	.31**	—	
3. RAPM	.31**	.42**	—
4. Letter sets	.41**	.13	.24**

Note: $N = 157$. WMC = working memory capacity composite score; gF = fluid intelligence; syllogisms = mean accuracy on the syllogistic reasoning task; RAPM = Raven Advanced Progressive Matrices.
**$p < .01$.

General discussion

Although there is a consistently observed relation between WMC and reasoning, the reasons behind this relationship are still not entirely clear. The present investigation examined belief bias, an element of reasoning that is the tendency to endorse conclusions as valid when they are believable and as invalid when they are unbelievable. More specifically, we wanted to see whether individual differences in WMC, which have been consistently shown to predict the ability to override automatic, pre-potent responses in favour of more controlled goal-directed responses, may shield individuals from belief bias.

We measured belief bias by giving participants four different types of syllogisms: valid with believable conclusions, valid with unbelievable conclusions, invalid with believable conclusions, and invalid with unbelievable conclusions. We replicated the belief bias effect in both accuracy and response times in two separate experiments with large sample sizes. Participants were less accurate when the validity and the believability of the conclusion conflicted, and they took longer to respond to these types of syllogisms. It was also clear that individuals employed controlled, System 2 processing to effectively reason in the presence of this conflict, as response times correlated with accuracy for conflict syllogisms, but not for no-conflict syllogisms. Although WMC offered a global benefit on the syllogistic reasoning task, WMC did not interact with the belief bias effect. In other words, participants with greater WMC were not less susceptible to belief bias. We also measured fluid intelligence, another aspect of reasoning, with two other tasks. A composite score from these tasks did not show an interaction with belief bias. In a follow-up experiment, we included decontextualized syllogisms to measure baseline logical reasoning abilities. We included these types of syllogisms to gain a better baseline from which to calculate the magnitude of the belief bias effect. Again, we found that the belief bias effect did not correlate with WMC (accuracy: $r = -.07$, $p = .38$; RT: $r = -.05$, $p = .55$) or fluid intelligence (accuracy: $r = .01$, $p = .92$; RT: $r = -.17$, $p = .05$).[7] These results were consistent for both accuracy and response times in both experiments.

These results are inconsistent with the findings of Stanovich and West (1998; see also Toplak, West, & Stanovich, 2011), who found that belief bias correlated with cognitive abilities, and with the findings of Quayle and Ball (2000), who found that individuals with greater spatial WMC were less susceptible to belief bias. However, Stanovich and West (1998) and Toplak et al. (2011) only gave participants conflict syllogisms. Therefore, the observed correlation could be due to a relationship between cognitive ability and baseline reasoning abilities. But in general, the results are consistent with the idea that cognitive abilities and cognitive biases are relatively independent sources of individual differences in reasoning (Stanovich & West, 2008). However, even when belief bias has been measured as a difference between accuracy on match syllogisms and conflict syllogisms, Sá et al. (1999) observed correlations with their cognitive ability measures, even though our observed average accuracy and differences were comparable. The results are also inconsistent with those of De Neys et al. (2005) and De Neys (2006). There are several possible reasons for this finding. In their experiments, De Neys et al. (2005) compared high- and low-WMC participants using a quartile split. De Neys (2006) used a full range of participants, but categorized participants as high-, mid-, and low-WMC. This procedure of turning a continuous variable into a categorical one can sometimes produce significant group differences that do not manifest when using WMC as a continuous variable (Rucker, McShane, & Preacher, 2015). Although we used this same procedure to illustrate the differences between high- and low-WMC participants, we used the full sample of participants and treated WMC as a continuous variable, using repeated measures ANCOVA for our analyses.

Another difference between the studies is the use of dual-task interference to test the hypothesis that high-WMC subjects are utilizing WMC to generate counter-examples during conditional reasoning. Although this technique has its merits, dual-task conditions often affect groups in different ways, and though this can allow inferences to be made about the processes used by groups in interference-free conditions, these effects are not always consistent. Therefore, we utilized a correlational approach, and this could account for the differences in findings between our study and that of De Neys et al. (2005). Additionally, there could be slight discrepancies between our study and those of De Neys et al. (2005) and De Neys (2006) in the way in which we measured WMC. De Neys et al. (2005) gave participants the operation span task in a group setting and selected high and low performers from a large group of participants. We gave participants the automated operation span task individually, as well as two other complex span tasks (i.e., reading span and symmetry span). However even comparing the upper and lower quartiles of participants on the automated operation did not reveal an interaction between this group difference and any of the observed main effects of validity or believability, nor the believability by validity interaction with either accuracy or response time. This was the case for both Experiment 1 and Experiment 2.

Alternatively, the differences in observations between De Neys et al. (2005) and the present study could be the different reasoning problems utilized. We attempted to give participants valid yet unbelievable conclusions (e.g., "All whales walk") and invalid yet believable conclusions (e.g., "All objects with four corners are rectangles") to force participants to override prior belief in the service of reasoning. De Neys et al. (2005) manipulated the interference of prior belief with the number of possible alternatives/disablers. In our view, this is the biggest difference between our study and that of De Neys et al. (2005). Our study is much more similar to that of De Neys (2006). The one major advantage our study has is the use of decontextualized syllogisms as an alternative way of measuring the magnitude of belief bias on an individual level. Although our study most directly conflicts with the findings of De Neys (2006), it is not entirely clear that the observed span X conflict interaction existed in the no load condition. All of these subtle differences could have caused the discrepancy in the findings, either individually or collectively. Because of the subtlety of the differences, it is hard to pinpoint the source of discrepancy. Further replications of these studies should provide a clearer answer.

These results cannot address the mental model theory of the relationship between working memory and reasoning (Copeland & Radvansky, 2004; Markovits et al., 2002), as we did not try to manipulate the number of mental models between syllogism type. Rather, our manipulations were only intended to vary the validity and believability of the conclusion. In the forms we used for the valid and invalid syllogisms, the invalid syllogisms did require more mental models, which was reflected in the main effect of validity for both accuracy and response time. However, all valid syllogisms and all invalid syllogisms had the same number of mental models, so we can still make comparisons between types of syllogisms with the same validity. The interaction between validity and believability, in which valid/unbelievable syllogisms showed lower accuracy and longer response times than valid/believable syllogisms, and invalid/believable syllogisms showed lower accuracy and longer response times than invalid/unbelievable syllogisms, reveals the effects of belief bias. As Stanovich and West (1998) theorize, the source of individual differences in reasoning can arise at several possible stages in the reasoning process. We attempted to isolate the differences to the stage at which individuals must sustain decontextualization of the information in order to reason independently of prior belief, which presumably requires some degree of cognitive capacity.

There are several possible explanations for why WMC was unrelated to belief bias in the present study. One possibility is that WMC is not a good candidate mechanism to explain the computational limitations of the sustained decoupling required for logical reasoning. However, because of the strong relationship between WMC and reasoning, especially at the latent level, we do not think this is the case. Another possibility is that resistance to belief bias is not a cognitive individual difference but a dispositional one. Stanovich and West (2008) show that cognitive abilities and thinking biases are relatively independent. In some cases they interact, and in some cases they do not. Therefore it is possible that a dispositional characteristic, such as need for cognition (see Cacioppo, Petty, Feinstein, & Jarvis, 1996), actively open-minded thinking, or cognitive reflection (Frederick, 2005; Toplak et al., 2011) better account for individual differences in belief bias. Although Stanovich and West (2008) argue that it is likely that individual differences will arise at the cognitive level when the task requires individuals to resolve conflict and compute an analytic response, that is not what we observe here. Rather, individual differences in WMC and gF seemed to offer a global benefit for those with better cognitive abilities.

Although we proposed WMC as a possible reason behind individual differences in resistance to belief bias because of the two-factor (i.e., goal maintenance and conflict resolution) theory of cognitive control (Engle & Kane, 2004), it could be that in this case conflict in verbal reasoning is qualitatively different from the types of conflict resolution individual differences in WMC predict. Recent work has attempted to narrow and define the boundary conditions for when WMC is an important individual difference and when it is not. For example Unsworth, Redick, Spillers, and Brewer (2012) found that although individual differences in WMC predicted goal maintenance on choice reaction time, anti-saccade, Stroop, and flanker tasks, high- and low-WMC participants did not differ in post-error slowing or conflict adaptation, so-called "micro-adjustments" of cognitive control. Meier and Kane (2015) found that WMC was related to stimulus–stimulus but not stimulus–response conflict. Therefore, WMC correlates with many, but not all, types of conflict resolution. In the present study, it could thus certainly be the case that resolving conflict in a verbal reasoning task is another type of conflict resolution that lies outside the boundary conditions of the WMC-cognitive control relation. From this perspective, the present results add to our ability to further delineate and define the precise nature of individual differences in WMC.

Conclusion

The present study sought to test the hypothesis that the sustained decoupling of information from its context is an individual difference reflected in working memory capacity (WMC). If greater WMC allows individuals to decontextualize information in the service of reasoning and resist cognitive biases, we should have seen a correlation between the magnitude of the belief bias effect (measured with accuracy and response time) and WMC. This correlation was not observed in either Experiment 1 or Experiment 2. Rather, WMC seems to offer a global benefit to reasoning. It seems to allow individuals to maintain and manipulate information in the service of reasoning, but it does not offer a specific resistance to belief bias. The results of the present study bring new evidence that WMC may not be the source of individual differences that allows some individuals to resist cognitive biases. The present study also narrows the scope of the predictive power of WMC. Recent work (e.g., Meier & Kane, 2015; Unsworth et al., 2012) has attempted to delineate and define the boundary conditions for WMC's role in human cognition. The present results are another step in helping to refine our understanding of individual differences in WMC.

Notes

1. An independent group of 28 participants rated the conclusions for degree of believability on a scale of 1–6 (1 = totally unbelievable, 6 = totally believable). A comparison of ratings for believable and unbelievable conclusions revealed that believable conclusions ($M = 5.46$, $SD = 0.62$) were rated as significantly more believable

than the unbelievable conclusions (*M* = 1.82, *SD* = 0.33); $t(27) = 26.74$, $p < .001$.
2. We repeated this analysis with a gF composite score (mean of standardized scores for Raven Advanced Progressive Matrices and letter sets) as a covariate, and we observed the same pattern of results.
3. We also ran the analyses on accurate RTs, and the pattern of results was identical, so we report RTs for all trials.
4. There were two participants with outlying data for the RT measures. Removing them from the analysis did not qualitatively change the pattern of results, so we included them in all analyses.
5. We again repeated these analyses with a gF factor score as a covariate, and the pattern of results was identical.
6. The degrees of freedom for this test are not a whole number because we used linear mixed modelling for this technique, as not all participants had an accurate and inaccurate response for each trial type.
7. A Bayesian approach to the correlations revealed Bayes factors all in favour of the null (WMC–accuracy = 6.25, WMC–RT = 7.69, gF–accuracy = 9.09, gF–RT = 1.35).

Disclosure statement

No potential conflict of interest was reported by the author.

References

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Copeland, D. E., & Radvansky, G. A. (2004). Working memory and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, 57A, 1437–1457.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433.

De Neys, W., Schaeken, W., & D'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking and Reasoning*, *11*, 349–381.

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

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

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*, 18423–18250.

Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes [Commentary on Stanovich and West]. *Behavioral and Brain Sciences*, *23*, 681–683.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, and Giroux.

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

Markovits, H., Doyon, C., & Simoneau, M. (2002). Individual differences in working memory and conditional reasoning with concrete and abstract content. *Thinking and Reasoning, 8*, 97–107.

Meier, M. E., & Kane, M. J. (2015). Carving executive control at its joints: Working memory capacity predicts stimulus-stimulus, but not stimulus-response, conflict. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *41*, 1849–1872.

Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, *53A*, 1202–1223.

Raven, J. C., Raven, J. E., & Court, J. H. (1998). *Progressive matrices*. Oxford, England: Oxford Psychologists Press.

Rucker, D. D., McShane, B. B., & Preacher, K. J. (2015). A researcher's guide to regression, discretization, and median splits of continuous variables. *Journal of Consumer Behavior*, *25*, 666–678.

Sá, W. C., West, R. F., & Stanovich, K. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, *91*, 497–510.

Stanovich, K. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Stanovich, K., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357.

Stanovich, K., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 161–188.

Stanovich, K., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Personality Processes and Individual Differences*, *94*, 672–695.

Thurstone, T. G. (1962). *Primary mental abilities*. Chicago: Science Research Associates.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*, 1275–1289.

Unsworth, N. (2014). Working memory capacity and reasoning. In A. Feeney, & V. Thompson (Eds.), *Reasoning as memory* (pp. 9–33). London: Psychology Press.

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.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J., & 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*, 635–654.

Unsworth, N., Redick, T. S., Spillers, G. J., & Brewer, G. A. (2012). Variation in working memory capacity and cognitive control: Goal maintenance and micro-adjustments of control. *Quarterly Journal of Experimental Psychology*, *65*, 326–355.