Mem Cogn
DOI 10.3758/s13421-014-0465-3

Individuals with low working memory spans show greater interference from irrelevant information because of poor source monitoring, not greater activation

Lindsey Lilienthal · Nathan S. Rose · Elaine Tamez · Joel Myerson · Sandra Hale

Abstract Although individuals with high and low working memory (WM) span appear to differ in the extent to which irrelevant information interferes with their performance on WM tasks, the locus of this interference is not clear. The present study investigated whether, when performing a WM task, high- and low-span individuals differ in the activation of formerly relevant, but now irrelevant items, and/or in their ability to correctly identify such irrelevant items. This was done in two experiments, both of which used modified complex WM span tasks. In Experiment 1, the span task included an embedded lexical decision task designed to obtain an implicit measure of the activation of both currently and formerly relevant items. In Experiment 2, the span task included an embedded recognition judgment task designed to obtain an explicit measure of both item and source recognition ability. The results of these experiments indicate that low-span individuals do not hold irrelevant information in a more active state in memory than high-span individuals, but rather that low-span individuals are significantly poorer at identifying such information as irrelevant at the time of retrieval. These results suggest that differences in the ability to monitor the source of information, rather than differences in the activation of irrelevant information, are the more important determinant of performance on WM tasks.

L. Lilienthal (✉) · E. Tamez · J. Myerson · S. Hale
Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA
e-mail: lindsey.lilienthal@yahoo.com

N. S. Rose
Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, USA

Keywords Working memory · Individual differences · Irrelevant information · Source memory

Working memory (WM), the ability to temporarily maintain and manipulate a limited amount of information (Baddeley & Hitch, 1974; Baddeley, 1986), has become an increasingly important construct in many areas of psychology. Individual differences in WM ability are important in a wide range of higher order cognitive functions, including reading comprehension (e.g., Daneman & Carpenter, 1980; Just & Carpenter, 1992), abstract reasoning (e.g., Engle, Tuholski, Laughlin, & Conway, 1999; Unsworth & Engle, 2007b), and learning (e.g., Lilienthal, Tamez, Myerson, & Hale, 2013; Kyllonen & Stephens, 1990). However, exactly what these individual differences in WM reflect is still somewhat unclear. Although numerous theories of WM posit possible sources of such individual differences, many of these suggestions lack sufficient empirical support.

One possible source of individual differences in WM is the ability to retrieve relevant information from secondary memory in the presence of irrelevant distractors, including the ability to resist proactive interference. Proactive interference occurs when previously acquired information disrupts one's ability to learn and remember new information (e.g., Keppel & Underwood, 1962), and the amount of proactive interference on a memory task has been shown to increase with the number of previous lists, negatively impacting memory performance. Many studies have shown that as the number of previous lists increases, participants tend to recall fewer correct items and are more likely to make an intrusion error, such as recalling an item from a previous list (e.g., Greenberg & Underwood, 1950; May, Hasher, & Kane, 1999; Kane & Engle, 2000).

The results of some studies have also suggested that individuals with low WM spans are significantly more affected by

Published online: 22 October 2014

proactive interference when performing a memory task than are individuals with high WM spans. For example, Kane and Engle (2000) had participants learn a number of lists of words from the same semantic category, and although the number of correct items that could be recalled decreased for everyone across subsequent lists, this decrease was significantly greater for those with low WM spans. In addition, Bunting (2006) directly manipulated proactive interference across trials of complex span task and found that only performance on high interference trials correlated with fluid intelligence. Moreover, several studies have reported that low-span individuals tend to make more intrusion errors than high-span individuals, and interpreted intrusions as reflecting susceptibility to proactive interference (e.g., Rosen & Engle, 1998; Unsworth, 2007; Unsworth & Brewer, 2010). However, most studies have not taken into account individual differences in the number of items recalled when comparing intrusion rates. That is, because low-span individuals, by definition, can recall fewer correct items, they may also be more likely to guess, which could result in them making more intrusion errors. In fact, when Rose (2013) controlled for differences in veridical recall on an immediate test, the difference in intrusion rate across WM span disappeared, suggesting that more research regarding intrusions may be needed.

It has been proposed that individual differences in the ability to resist the effects of PI reflect differences in inhibition or executive attention, which are involved in restricting one's attention to only relevant information, with high-span individuals better able to restrict their attention (e.g., Hasher & Zacks, 1988; Engle, 2002). This suggests that high-span individuals should show the greatest level of activation for memory items from the current memory list, whereas low-span individuals, who have more difficulty limiting their attention, may show similar levels of activation for memory items from the current memory list and for memory items from previous lists.

However, no study has directly tested whether high- and low-span individuals truly differ in the activation strength of formerly relevant, but now irrelevant, items from previous lists on WM tasks. To the best of our knowledge, Experiment 1 of the present study represents the first direct examination of whether high- and low-span individuals differ in such activation when performing a WM task.

Participants in Experiment 1 performed an operation span task, but following some series of to-be-remembered words, rather than being asked to recall the words, participants instead were asked to perform a lexical decision task (Meyer & Schvaneveldt, 1971). Participants respond more quickly to items that are being held in a more active state, and thus lexical decision tasks provide implicit measures of the activation levels of items or concepts in memory (e.g., Ratcliff, Hockley, & McKoon, 1985).

Critically, some of the words used in the lexical decision task had previously appeared as to-be-remembered words in the operation span task, some on the current list and others on the previous list. If low-span individuals hold more irrelevant information in an active state than high-span individuals, then high-span individuals should be slower to respond to words from the previous list (i.e., formerly relevant, but now irrelevant words) than to words from the current list (i.e., currently relevant words), presumably because they have more successfully limited their attention and/or inhibited previous information, whereas the response times (RTs) of low-span individuals to previous memory words should be more similar to their RTs to the current memory words, presumably because the previous words are still in an active state. The goal of the first experiment of the present study was to test this prediction regarding individual differences in activation strength on a WM task.

Importantly, some research suggests that another possible source of individual differences in WM is the ability to monitor the source of remembered information (e.g., Unsworth & Brewer, 2009; Unsworth & Brewer, 2010). Consistent with this interpretation, Rose (2013) found that when participants performed a complex span task and then were asked to recall both memory items and secondary task items on a surprise delayed test, and also to report the source of each item they recalled, low-span individuals made significantly more source errors than high-span individuals.

It is thus possible that low-span individuals are less accurate than high-span individuals when recalling on WM tasks not because they are more likely to have irrelevant information active, but because when they retrieve such information, they are better able to identify and exclude those items from their response. If that is the case, then one would expect to find differences in the ability to explicitly judge the source of a word during a complex span task. Although there is evidence suggesting that low-span individuals may have poorer source memory, no past research has investigated whether low-span individuals have poorer source memory in the context of a WM task, where source refers to whether the item came from the current list of to-be-remembered items or from a previous, now irrelevant list, and the goal of Experiment 2 of the present study was to examine this possibility.

Experiment 2 was designed to be analogous to Experiment 1, except that in the second experiment, recognition and source judgment tasks were used in place of the lexical decision task in the first experiment. More specifically, participants performed an operation span task, but following some series of to-be-remembered words, rather than recall the words, participants were asked to perform recognition and source judgment tasks. Participants were asked to respond 'old' to any word that had appeared as a to-be-remembered word in the task, and after every 'old' response, participants were asked to specify whether the word had been presented as part of the current memory list or the previous list. If low-span individuals have more difficulty identifying whether an item was presented as part of the most recent series or the previous series than high-span individuals, then differences in source judgment accuracy should be observed.

Experiment 1

The first experiment was designed to investigate whether the activation of formerly relevant, but now irrelevant items differs as a function of WM span. The procedure used was similar to an operation span task (Turner & Engle, 1989), but with one critical manipulation (see Fig. 1). As in an operation span task, on every trial participants were asked to encode a series of to-be-remembered words, and to perform a math operation task following the presentation of each word. However, participants were only asked to recall the to-be-remembered words on half of the trials. On the other half of trials (intermixed in a pseudorandom order), after encoding the to-be-remembered words and performing the math operations, participants performed a lexical decision task in which some of the words had appeared as to-be-remembered items on either the current trial or the previous trial. RTs to these current and previous words provided estimates of relative state of activation for both types of memory items. If low-span individuals have more irrelevant information activated in memory as compared to high-span individuals, we would expect to see a significant interaction between WM span and word type.

Method

Participants

All 83 participants (58 female; mean age =19.2, *SD* =1.3) were Washington University undergraduates who participated in exchange for course credit, and who reported English as their native language.

Materials and procedure

Participants were tested individually during a single experimental session that lasted approximately 1.5 h. All participants completed a reading span task in addition to the operation span/lexical decision task. Participants also completed the Ravens Advanced Progressive Matrices (RAPM; Raven, Raven, & Court, 1998) and a category-judgment RT task (for details, see Chen, Myerson, & Hale, 2007), which were originally included for another purpose irrelevant to the present experiments and therefore will not be reported here.

Operation span/lexical decision task A schematic of the procedure for this task is presented in Fig. 1. Participants saw a series of five words, each presented on the screen for 1500 ms.

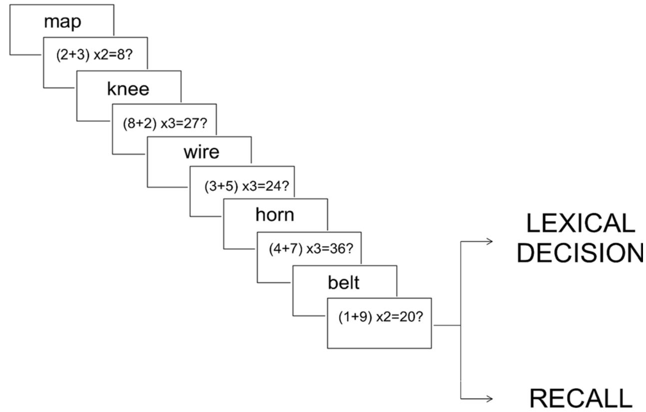

Fig. 1 Schematic representation of the operation span/lexical decision task used in Experiment 1. Half of the trials ended in recall, the other half ended in a lexical decision task

Following the presentation of each word, participants were shown a math problem (e.g., (3 +6) ×2 =18?) and were asked to decide whether or not the given solution was correct. Participants indicated their decision by pressing one of two keys on the computer keyboard. Each math problem remained on the screen until a response was made or until an individually set time limit had expired (more detail about this time limit is provided in the description of the first round of practice).

Importantly, following each series of words, participants were asked to do one of two things. On half of the trials, participants were asked to recall the preceding five words in the correct serial order, and to do so by typing in the words. On the other half of trials (intermixed in a pseudorandom order), participants were shown a series of letter strings and were asked to decide whether or not each letter string constituted a real English word. Participants responded to the letter strings using the same two keys that were used to respond to the math problems, and were instructed to respond as quickly and accurately as possible.

Participants did not know whether recall or lexical decisions would be required on each trial until after the five to-be-remembered words were presented. Pilot tests indicated that RTs on the first few letter strings presented during the lexical decision task were much slower than on the rest of the strings. Accordingly, three buffer letter strings (non-words and new words, none of which had been presented previously in the task) were included to allow participants to become accustomed to responding to the lexical decision task. Responses on these buffer letter strings were not included in analyses. Following the three buffer letter strings, participants were presented with either a non-word or a new word, followed by a word that had appeared as a to-be-remembered item. On a pseudorandom half of these lexical decision trials, this was a current memory word (i.e., a word that had been presented as part of the current list of five to-be-remembered words), and on the other half of these lexical decision trials, this was an old memory word (i.e., a word that had been presented as part of the previous list of five to-be-remembered words). After responding to the memory word from the current or previous list, participants were presented with three additional letter strings, two non-words, and one new word. Thus, on each trial with lexical decisions, participants were presented with a total of eight letter strings, four of which were real words.

Participants were not informed that memory words would be included as a part of the lexical decision task. RTs on these words provided a measure of the activation strength of currently relevant words as well as formerly relevant, but now irrelevant words. The RT measures for new words and non-words were provided by the responses to the letter string immediately preceding the memory word as well as the three letter strings that followed the memory word. After the recall or lexical decision task, participants were given feedback regarding their accuracy on the math problems (i.e., participants were presented with a percentage reflecting the number of the preceding five problems answered correctly) in order to encourage proper completion of this portion of the WM task.

Participants began the operation span/lexical decision task by completing four rounds of practice. During the first round, they were presented with a series of 30 math problems and were asked to respond as to whether or not each solution was correct. RTs during this first round of practice were used to set individual time limits for subsequent math problems on both practice and test trials so as to minimize the possibility that participants would slow down on the math problems in order to covertly retrieve the to-be-remembered words. Specifically, the time limit was set equal to a participant's mean RT plus two standard deviations, and if a response did not occur before the time limit expired, the computer program moved on and the math problem was scored as incorrect.

During the second round of practice, participants were presented with two series of five to-be-remembered words which they were asked to recall; this second round of practice was to familiarize participants with the recall procedure and did not include math problems. During the third round of practice, participants completed three recall trials that were identical to the recall test trials (i.e., a math problem was presented following each memory word). For the fourth round of practice, participants received instructions for trials with lexical decisions (instead of recall) and they completed three such trials that were identical to the lexical decision test trials except that no memory words were included.

Following the fourth round of practice, participants completed a total of 60 test trials with a 30-s break after every 12 trials. As mentioned, half of the trials involved recall and the other half involved lexical decisions. Importantly, participants had no way of knowing whether a trial would involve recall or lexical decisions until after they had been presented with all five words and math problems, and thus they were told to always try to remember the words. The two trial types alternated in a way that appeared to be random but that was the same for every participant. The very first test trial, as well as the first trial following each 30-s break, always ended in recall.

Reading span task In this verbal complex span task (Daneman & Carpenter, 1980), participants saw a series of single digits, each presented on the computer screen for 1500 ms. Preceding each digit, participants were shown a sentence that they were asked to read aloud and then decide whether or not it was meaningful. Sentences that were not meaningful were created by replacing one word in an otherwise meaningful sentence (e.g., "My mother always told me it is not hungry to boast."). Participants indicated their decision by pressing one of two keys on the computer keyboard; if a response was not made within 10 s, the program moved on to the next to-be-remembered digit and the decision was scored as incorrect. At the end of each series, participants were asked to recall the digits aloud in the order of their presentation, and these vocal responses were recorded using an Olympus VN-3100PC digital recorder.

Participants began the task with two rounds of practice: During the first round, participants responded to four sentences in the absence of any memory task, and during the second round, they completed three trials identical in procedure to the test trials. Participants then completed two test trials at each list length from two to seven, for a total of 12 test trials. Performance on the reading span test trials, assessed using a partial scoring method in which participants were awarded one point for each digit recalled in the correct serial position, provided a measure of WM span separate from performance on the operation span/lexical decision task.

Analysis

To avoid the problems associated with dividing a continuous variable like WM span into discrete groups, the RT data for all four word types were entered into a repeated-measures generalized linear model (GLM) analysis with WM (i.e., reading span task performance) included as a covariate in order to determine whether the degree of activation changed as a function of WM span. If the degree of activation did change as a function of span, this would be revealed by a significant interaction between word type and WM. The data from four participants demonstrated significant Mahalanobis Distance values indicating that they were multivariate outliers, and thus these data were excluded and analyses were conducted on data from 79 participants.

Results and discussion

Descriptive statistics for Experiment 1 are provided in Table 1. The repeated-measures GLM analysis, which was the primary focus of the current experiment, revealed effects of both word type, $F(3, 231) = 4.01$, $p = .008$, and WM span, $F(1, 77) = 5.35$, $p = .02$, on RT, but the interaction that would have been expected if low-span individuals have more irrelevant information activated in memory was not significant, $F(3, 231) = 0.57$, $p = .63$. That is, the absence of an interaction implies that individuals with higher and lower WM spans did not differ in the degree to which formerly relevant, but now irrelevant information was activated compared to currently relevant information. This is illustrated in the top panel of Fig. 2, which plots the difference between RTs for old and current words as a function of WM span.

In addition, the absence of an interaction was confirmed by a planned comparison using GLM that revealed no significant interaction between word type and WM span when the analysis was restricted to current and old words, $F(1, 77) = 0.07$. Although the inhibitory and controlled attention views of differences in span are somewhat unclear as to when inhibition and controlled attention operate to reduce PI on complex span tasks, if the activation of now-irrelevant items were suppressed at the beginning of each new trial, and low-span individuals were less able to suppress items from the preceding list (i.e., old words), then they should have been faster to respond to old words relative to current words, and thus an interaction should have been observed. Alternatively, if the activation of now-irrelevant items were suppressed at the beginning of the lexical decision task, then the proper comparison would be between RTs to current words and new words. That is, if low-span individuals were less able to suppress current words, then such individuals should be faster to respond to current words relative to new words than higher-span individuals: When the GLM analysis was restricted to current and new words, however, again no significant interaction between word type and WM span was observed, $F(1, 77) = 0.70$.

Regarding the main effect of word type, further planned comparisons revealed that old word RTs were significantly slower than current word RTs, $t(78) = 2.04$, $p = .04$, new word RTs were significantly slower than old word RTs, $t(78) = 2.01$, $p < .04$, and nonword RTs were significantly slower than new word RTs, $t(78) = 7.51$, $p < .001$ (see the top, middle, and bottom panels of Fig. 2, respectively). Thus, priming was significantly greater for current words than for formerly relevant, but now irrelevant words.

The finding that participants responded to both current and old memory words faster than they responded to new words indicates that both types of memory words were significantly primed compared to baseline. Critical to our research question, however, was the interaction between word type and WM span. If low-span individuals have more irrelevant information active following the presentation of a memory list than high-span individuals, due to a greater susceptibility to proactive interference, one would expect that the activation of formerly relevant, but now irrelevant memory words would depend on one's span. However, this was not the case, as the degree of priming did not change as a function of WM span, implying that individuals with lower WM spans do not maintain representations of formerly relevant, but now irrelevant words from past memory lists in a more active state than individuals with higher WM spans.

Table 1 Descriptive statistics for Experiment 1

Measure	*M*	*SD*	Range
Reading Span			
Recall	.68	.14	.41 – 1.0
Sentence Accuracy	.96	.02	.89 – 1.0
Operation Span/Lexical Decision			
Recall	.71	.19	.20 – .99
Math Accuracy	.93	.04	.79 – .99
Current Word RT	579	72	446 – 792
Old Word RT	590	80	460 – 843
New Word RT	601	65	479 – 872
Non-word RT	641	69	516 – 858
Current Word Accuracy	.99	.02	.87 – 1.0
Old Word Accuracy	.99	.03	.80 – 1.0
New Word Accuracy	.99	.03	.91 – 1.0
Non-word Accuracy	.98	.04	.85 – 1.0

Note. Current Word RT, Old Word RT, New Word RT, and Non-word RT data are reported in milliseconds. Data from all other measures are reported as the proportion of items correct. RT = response time

Experiment 2

The second experiment was designed to investigate whether high- and low-span individuals differ in their ability to correctly identify the source of a recently presented memory item on a WM task, specifically whether the item was presented as part of the current list of to-be-remembered words or the previous list. The results of Experiment 1 suggest that high- and low-span individuals do not differ in the activation of formerly relevant, but now irrelevant memory items on WM tasks, but it is possible that they may differ in their ability to use list-level context cues in order to correctly identify the source of memory items. Experiment 2 was designed to test this hypothesis, and as in Experiment 1, this was done using a modified operation span task.

As in Experiment 1, on each trial of the modified operation span task participants were asked to encode a series of to-be-remembered words and to perform a math operation task

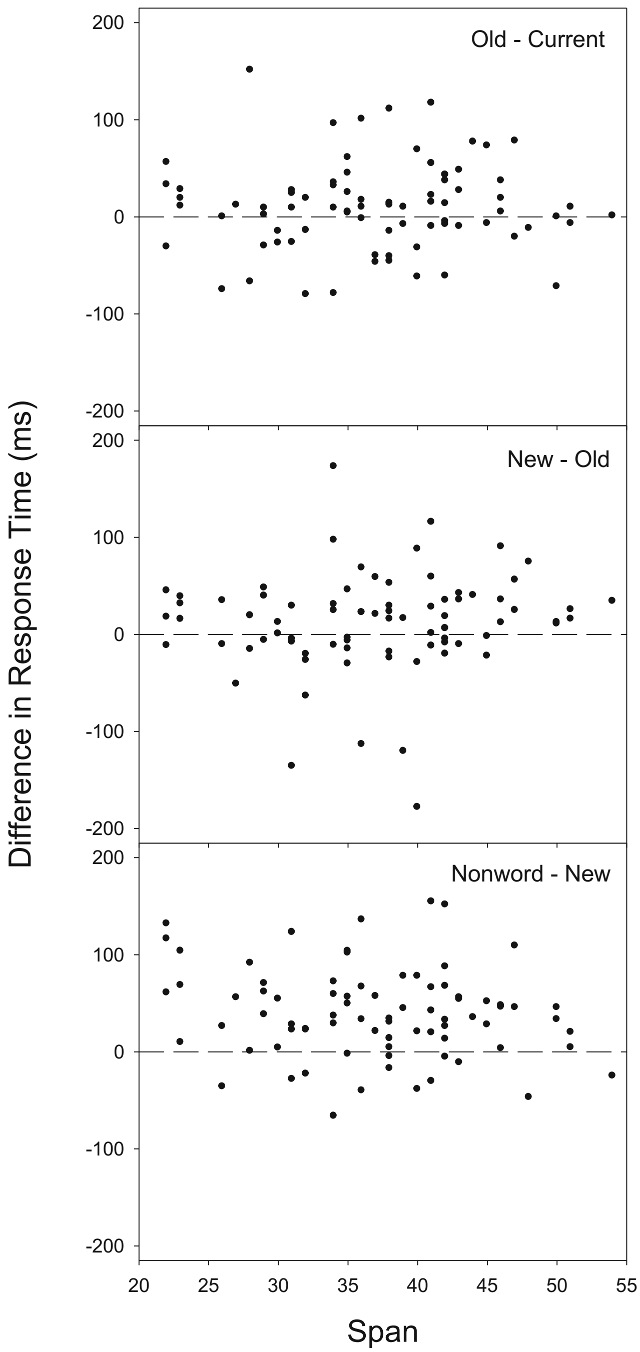

Fig. 2 Differences in response time (RT) between old and current words (top panel), new and old words (middle panel), and non-word and new words (bottom panel) as a function of working memory (WM) span in Experiment 1

following the presentation of each word. Critically, however, participants were only asked to recall the to-be-remembered words on half of the trials. On the other half of trials (intermixed in a pseudorandom order), participants instead performed an item recognition task in which they saw a series of words and were asked to indicate whether they were old or new, that is, whether or not they had been presented at any time during the modified operation span task. Following every 'old' response, participants were asked to make a source judgment, indicating whether the word had been presented as part of the current list or the previous list.

Method

Participants

All 88 participants (45 female; mean age =19.3, *SD* =1.1) were Washington University undergraduates who participated in exchange for course credit, and who reported English as their native language.

Materials and procedure

All participants in this experiment completed a standard operation span task, followed by an operation span/recognition task. Participants also completed the odd-numbered items from the RAPM for a purpose irrelevant to the present experiments and therefore will not be reported here. All participants were tested individually, and both tasks were completed during a single experimental session lasting approximately 1.5 h.

Operation span/recognition task A schematic of the procedure for this task is presented in Fig. 3. Just as in the operation span/lexical decision task used in Experiment 1, participants were asked to remember a series of five words, perform a math operation following the presentation of each word, and, on half of the trials, to recall the words in serial order. The procedure on these recall trials was identical to the corresponding trials of the operation span/lexical decision task in Experiment 1.

On the other half of the trials (intermixed in a pseudorandom order), rather than recalling the five words that had just been presented, participants instead were asked to perform a recognition task in which they were shown a series of words one at a time and asked to decide whether each was a new word (i.e., a word that had not been presented as a memory word during the task) or an old word (i.e., a memory word). Participants indicated this decision by pressing one of two keys on the computer keyboard. Importantly, participants were instructed to respond 'old' to any word that had been presented as part of the modified operation span task, regardless of whether or not it was from the current list of five to-be-remembered words. Following every 'old' response, participants were asked to report the source of the 'old' word, that is, whether it came from the current list or from the previous list, again by pressing one of two keys on the computer keyboard.

On each trial with a recognition task, participants were presented with a total of 13 words. The first presented word was always a new word that served as a buffer word and was

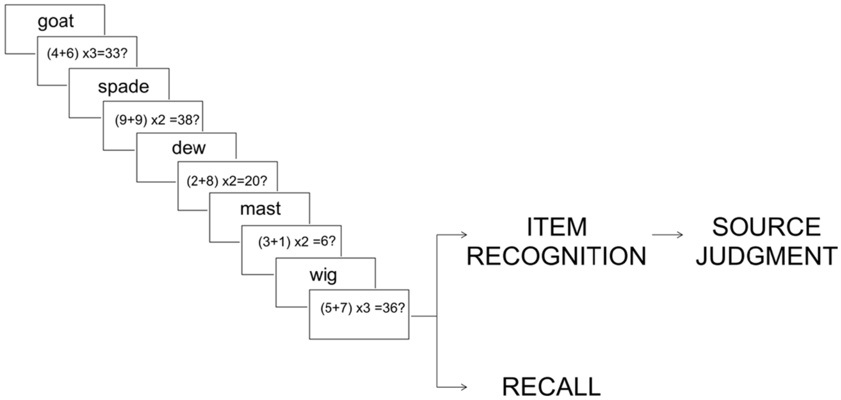

Fig. 3 Schematic representation of the operation span/recognition task used in Experiment 2. Half of the trials ended in recall, the other half ended in an item recognition and source judgment task

not included in analyses. Following the initial buffer word, participants were presented with six new words and six old words in a random order. Half of the old words were from the current list and half were from the previous list. As already mentioned, participants were also asked to make a source judgment for every word to which they responded 'old'. This was the case even if the 'old' response was an error (i.e., even if the word had actually been new), although source judgments to new words (which constituted 5 % of the words that participants identified as old) were not included in analyses. Following the recall or recognition task on each trial, participants were given feedback regarding their accuracy on the preceding five math problems.

Participants completed four rounds of practice. The first three rounds were identical to those used in Experiment 1. During the fourth round of practice, participants received instructions for trials that included a recognition task (instead of recall) and completed three such trials.

Following the fourth round of practice, participants completed a total of 60 test trials with a 30-s break after every 12 trials. As mentioned, half of the trials involved a recall task and the other half involved a recognition task Importantly, participants had no way of knowing whether a trial would involve recall or recognition until after they had been presented with all five words and math problems, and thus they were told to always try to remember the words. The two trial types alternated in a way that appeared to be random but that was the same for every participant. The very first test trial, as well as the first trial following each 30-s break, always ended in recall.

Operation span task In this verbal complex span task (Turner & Engle, 1989), participants were presented with a series of words, each of which was presented on the computer screen for 1500 ms. Preceding each word, participants were shown a math problem (e.g., (4 ×2) – 1 =7?) that they were asked to read aloud and then decide whether or not the solution given was correct. Participants indicated this decision by pressing one of two keys on the computer keyboard; if a response was not made within 10 s, the program moved on to the next to-be-remembered word and the decision was counted as incorrect. At the end of each trial, participants were asked to recall the words aloud in the order of their presentation, and these vocal responses were recorded using an Olympus VN-3100PC digital recorder.

Participants began the task with two rounds of practice: During the first round, participants responded to six math problems in the absence of any memory task, and during the second round, they completed four trials identical to the test trials. Participants then completed two test trials at each list length from two to seven, for a total of 12 test trials. Performance on the operation span test trials, assessed using a partial scoring method in which participants were awarded one point for each word recalled in the correct serial position, provided a measure of WM span separate from performance on the operation span/recognition task.

Analysis

D-prime values were calculated for both the item recognition judgments and the source recognition judgments on the operation span/recognition task. Accuracy was relatively high on this task, and some participants had hit rates of 1.00 or false alarm rates of 0.00; in such cases, d-prime was calculated using the correction suggested by MacMillan and Creelman (2005). In order to determine whether the ability to correctly recognize memory items and the ability to correctly judge their source vary as a function of WM span, d-prime values for both item recognition and source recognition were entered into a repeated-measures GLM analysis with WM (i.e., operation span task performance) included as a covariate. Critical to the hypothesis was the interaction between recognition type (item versus source) and WM span.

The data from seven participants were excluded from analyses; two were excluded due to significant Mahalanobis Distance values indicating that they were multivariate outliers, and five were excluded because their accuracy on the math problems of either the operation span/recognition task or the operation span task was below 75 %, and thus analyses were conducted on data from 81 participants.

Results and discussion

Descriptive statistics for Experiment 2 are provided in Table 2. The repeated-measures GLM analysis revealed effects of both recognition type (item versus source), $F(1, 79)$ =22.71, $p <$ .001, and WM span, $F(1, 79)$ =12.08, p = .001, on d-prime, as well as a significant interaction between recognition type and WM span, $F(1, 79)$ =8.54, p = .005. This interaction, which may be seen in Fig. 4, reflects the fact that although both the item recognition d-prime and the source recognition d-prime were significantly correlated with WM span, the correlation between source recognition d-prime and WM span (r = .41) was significantly greater than the correlation between item recognition d-prime and WM span, (r = .25), Z = −2.11, p = .035 (Lee & Preacher, 2013).

To explicate the observed differences in d-prime, hits and false alarms for both item recognition and source recognition were also considered separately (see Table 3). For item recognition, WM span was significantly related to hit rates but not to false alarm rates, suggesting that when determining whether an item had been seen previously, high-span individuals were more likely to correctly respond 'old' to previous items but no less likely to respond 'old' to a new item compared to low-span individuals. For source recognition, WM span was significantly related to both hit and false alarm rates, suggesting that when determining whether an item had been presented as a part of the current list or the previous list, high-span individuals were more likely to correctly respond 'current list' to items from the current list and less likely to respond 'current list' to items from the previous list compared to low-span individuals. Thus, the relation between WM span and d-prime seems to be driven primarily by hit rates for item recognition but by both hit and false alarm rates for source recognition.

The results of the present experiment are consistent with those of previous studies (e.g., Rose, 2013; Unsworth & Brewer, 2009; Unsworth & Brewer, 2010; Oberauer, 2005) in suggesting that, in addition to having poorer recognition memory overall, low-span individuals appear to have a particular deficit in the ability to identify the source of recently presented memory words. However, the present study is the first to show that high- and low-span individuals differ in their ability to judge whether memory words were presented as part of the current list or a previous list on a WM complex span task.

Table 2 Descriptive statistics for Experiment 2

Measure	*M*	*SD*	Range
Operation Span/Recognition			
Recall	.52	.28	.05 – .97
Math Accuracy	.94	.04	.75 – 1.0
Item Hits	.84	.09	.51 – 1.0
Item False Alarms	.05	.06	.00 – .41
Source Hits	.89	.08	.67 – 1.0
Source False Alarms	.15	.11	.00 – 43
Operation Span			
Recall	.45	.17	.07 – .81
Math Accuracy	.92	.05	.79 – .98

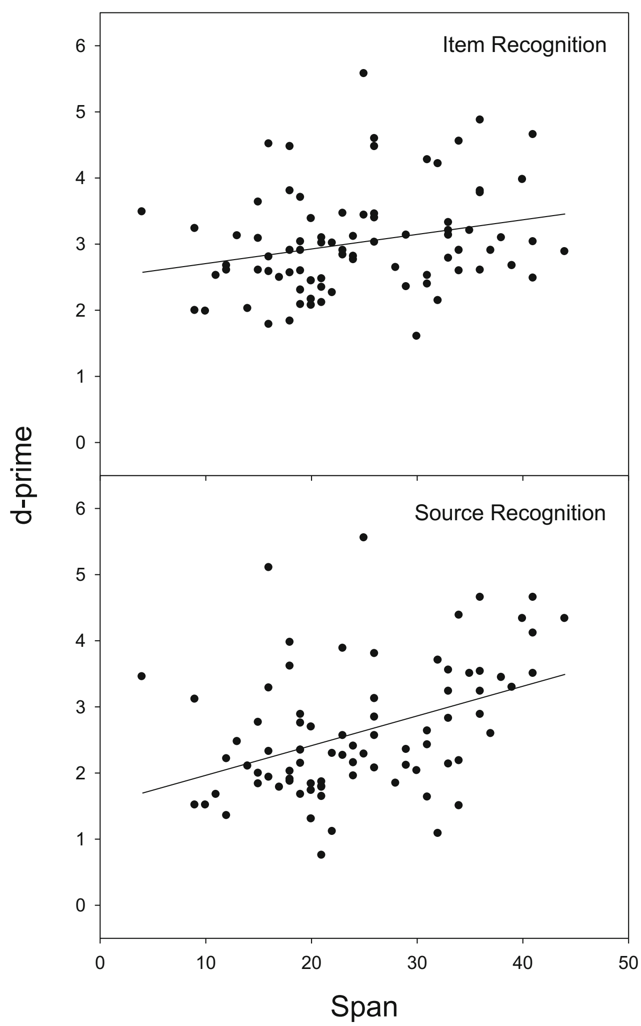

Fig. 4 D-prime values for item recognition and source recognition as a function of working memory (WM) span in Experiment 2

General discussion

It has been hypothesized that individual differences in WM span reflect differences in the extent to which individuals are affected by proactive interference, or, more generally, the extent to which irrelevant information interferes with their ability to recall (e.g., Hasher, Lustig, & Zacks, 2007; Engle, 2002; Unsworth & Engle, 2007a). Taken together, the results

Table 3 Correlations for Experiment 2

Measure	WM Span	Hits	False Alarms
WM Span	-	**.32**	-.06
Hits	**.33**	-	.09
False Alarms	**-.32**	**-.30**	-

Note. Correlations presented above the diagonal refer to item recognition, whereas correlations presented below the diagonal refer to source recognition. Significant correlations are in bold (all $p < .01$). WM =working memory

of the present two experiments suggest that when performing WM tasks, low-span individuals differ from high-span individuals not in the activation of currently and previously relevant items but in their ability to distinguish between the two.

Experiment 1 addressed the question of whether high- and low-span individuals differ in the activation strength of formerly relevant, but now irrelevant words using an implicit measure of activation. This was done using a modified operation span task on which half of the trials ended with lexical decisions rather than recall. In addition to non-words and new words, the stimuli for the embedded lexical decision task included words that had been presented as part of the current list as well as words that had been presented as part of the previous list, thereby making it possible to compare the degree of activation of both relevant and irrelevant words. If items that were recently relevant, but are now irrelevant, remain in a more active state in memory for low-span individuals due to their poorer inhibition or attentional control, then lexical decision RTs should have shown a significant interaction between WM span and word type.

Contrary to this prediction, however, the pattern of RTs across word type did not differ based on individuals' WM span. That is, both high- and low-span individuals showed elevated activation for both currently relevant and formerly relevant, but now irrelevant memory words compared to words that had not been presented previously as part of the task. This result is consistent with Unsworth and Engle (2005), who found that high- and low-span individuals did not differ in their implicit awareness on a serial reaction time task under incidental learning conditions. However, the present study is the first to directly test the implicit activation of current and previous items on a WM span task.

Experiment 2 addressed the question of whether high- and low-span individuals differ in their ability to explicitly judge the source of memory words and determine whether the words had been presented as part of the current list or the previous list. This was done using a procedure analogous to that used in Experiment 1: Item recognition and source judgments were embedded in a modified operation span task in Experiment 2 just as lexical decisions had been embedded in an operation span task in Experiment 1. During the embedded item recognition task, participants were asked to distinguish between new words (i.e., words that had never been presented as a memory word) and old words (i.e., memory words), and then report whether 'old' words were from the current or the previous list. If low-span individuals have a specific problem discriminating current memory items from formerly relevant, but now irrelevant memory items, there should have been a significant interaction between WM span and type (item versus source) of recognition judgment.

The results of Experiment 2 were consistent with this prediction: Although low-span individuals were less accurate at both types of recognition judgments, they were particularly poor at judging whether a word had been part of the current list or a previous list. This result is consistent with previous findings suggesting that individual differences in WM may be related to source memory, but these previous studies largely investigated source memory on long-term (episodic) memory tasks rather than on WM tasks: For example, both Unsworth & Brewer (2010) and Rose (2013) showed a relation between WM and individual differences in source memory for items on delayed free recall tests. Thus, this is the first study that has directly compared source memory for currently relevant items and previously relevant, but currently irrelevant, items from an ongoing WM task. It is frequently hypothesized that identical secondary/long-term memory components are involved in supraspan episodic memory tasks and WM tasks, and the present results provide support for this hypothesis.

Taken together, the present findings suggest that although high- and low-span individuals have equivalent *implicit* access to currently and formerly relevant items from WM tasks, low-span individuals have significantly poorer *explicit* access to such items, especially with regard to the temporal-contextual details on which source judgments are based. Thus, as has been previously hypothesized (e.g., Hasher et al., 2007; Engle, 2002; Unsworth & Engle, 2007a), the degree to which irrelevant information interferes with retrieving items from secondary memory plays an important role in individual differences in WM span. However, the locus of the influence of irrelevant or distracting information was not previously clear. At least with respect to formerly relevant, but now irrelevant items on WM tasks, the present findings suggest that high- and low-span individuals' differential susceptibility to interference occurs not because of differences in the activation level of irrelevant items, but rather because of differences in the ability to discriminate relevant and irrelevant items at retrieval.

Previous studies have shown that low-span individuals are more susceptible to proactive interference, as evidenced by low-span individuals recalling significantly fewer items than high-span individuals under conditions of high proactive interference (e.g., Kane & Engle, 2000) and by low-span individuals making significantly more intrusion errors (e.g., Rosen & Engle, 1998; Unsworth, 2007). Previous studies have also suggested that low-span individuals may have more difficulty monitoring the source of information, and may not recognize when retrieved information is irrelevant (e.g., Unsworth & Brewer, 2009; Unsworth & Brewer, 2010). However, most of this past research investigated possible differences between high- and low-span individuals on episodic memory tasks rather than WM tasks. Because the present experiments examined the role of irrelevant information in differences in WM span in a new, more direct way, the results clarify the association between individual differences in WM and output monitoring processes. Most importantly, the results of the present experiments suggest that the ability to monitor the source of information, and not the ability to inhibit or suppress irrelevant information, was the important determinant of individual differences in performance on WM tasks.

Acknowledgments The efforts of Lindsey Lilienthal were supported by NIA Training Grant AG00030.

References

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Bunting, M. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 183–196.

Chen, J., Hale, S., & Myerson, J. (2007). Predicting the size of individual and group differences on speeded cognitive tasks. *Psychonomic Bulletin & Review, 14,* 534–541.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23.

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.

Greenberg, R., & Underwood, B. J. (1950). Retention as a function of stage of practice. *Journal of Experimental Psychology, 40,* 452–457.

Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. Miyake, A. R. A. Conway, C. Jarrold, M. J. Kane, & J. N. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York: Oxford University Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). New York: Academic Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 336–358.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153–161.

Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2,* 129–160.

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org

Lilienthal, L., Tamez, E., Myerson, J., & Hale, S. (2013). Predicting performance on the Raven's Matrices: The roles of associative learning and retrieval efficiency. *Journal of Cognitive Psychology, 25,* 704–716.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (2nd ed.). Mahwah, N.J.: Lawrence Erlbaum Associates.

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

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term retention. *Journal of Experimental Psychology: General, 134,* 368–387.

Ratcliff, R., Hockley, W., & McKoon, G. (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General, 114,* 435–450.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.

Rose, N. S. (2013). Individual differences in working memory, secondary memory, and fluid intelligence: Evidence from the levels-of-processing span task. *Canadian Journal of Experimental Psychology, 67,* 260–270.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39,* 418–436.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154.

Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 1020–1034.

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,* 1578–1585.

Unsworth, N., & Brewer, G. A. (2010). Variation in working memory capacity and intrusions: Differences in generation or editing? *European Journal of Cognitive Psychology, 22,* 990–1000.

Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory and Cognition, 33,* 213–220.

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,* 104–132.

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,* 1038–1066.