Individual differences in working memory capacity affect false memories for missing aspects of events

Matthew P. Gerrie and Maryanne Garry

Victoria University of Wellington, New Zealand

Recently, Gerrie, Belcher, and Garry (2006) found that, when participants watch an event with parts missing, they falsely claim to have seen the missing parts—but they were more likely to claim they had seen less crucial parts than more crucial parts. Their results fit with a source-monitoring framework (SMF; Johnson, Hashtroudi, & Lindsay, 1993) explanation of false memories. In this paper we used individual differences in working memory capacity (WMC) to examine the SMF explanation of false memories for missing aspects of events. An accumulating body of research suggests that WMC is strongly related to controlling attention, including the ability to distinguish between sources of information. The primary purpose of the present study was to examine whether people with larger WMC are better able to ward off false memories for missing information than those with a smaller WMC. We showed that higher WMC reduced false recognition of crucial information, but did not change false recognition of noncrucial information. Additionally, we found that WMC had little effect on participants’ subjective experience of true and false recognition of events, regardless whether the information was crucial or not. These results provide further evidence that people’s WMC is related to their source-monitoring ability.

Although most false memory research has focused on the effect of explicit suggestions, there is a body of research that focuses on false memories that occur in the absence of explicit suggestion (Jenkins, Wald, & Pittenger, 1986; Reinitz, Lammers, & Cochran, 1992). Of this second line of research, the Deese (1959), Roediger and McDermott (1995; DRM) paradigm is the best known. In the DRM, participants are presented with a list of semantically related words (e.g., bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn, drowsy) that are all highly associated with an unpresented lure word (e.g., sleep). Later, participants correctly remember many of the words that were actually on the list (old words); correctly reject unrelated unseen words (new words); and often systematically—but falsely—remember the unseen lure word (critical lures).

The DRM could be criticised for not being very event-like, in that there is no sequence of actions that unfold over time. But other researchers have found systematic false memories with more event-like materials (Jenkins et al., 1986; Shaw, Wilson, & Wellman, 1986). For example, Bransford and Franks (1971) asked participants to read sentences describing events of varying complexity, and found that participants falsely recognised new sentences that were grafted together from smaller sentences. More recently, Chan and McDermott (2006) found that when participants read sentences describing an event,
they often falsely recognised new sentences that went beyond the meaning of the original. Still other research has used visual materials to depict longer event sequences. Hannigan and Reinitz (2001) showed participants slide sequences of everyday events (such as a woman grocery shopping) and found that when participants saw a slide of the woman standing next to some oranges on the floor, they falsely recognised having seen her remove an orange from a stack. Taken together, these studies show that people can come to falsely recognise an unseen slide or sentence from an event. In these studies, people are falsely remembering a single instance in time: a word from a list, a sentence from a narrative, or a slide from a sequence.

Recently, however, Gerrie et al. (2006) showed that people could falsely recognise entire chunks of a movie they have never seen—in effect, falsely remembering entire chunks of time. They first asked independent observers to watch a movie of a woman making a sandwich, and classified each action in the movie as either crucial or not crucial to the overall event. Then, in one study, when participants watched an event riddled with missing actions, they falsely claimed to have seen the missing actions—but they were more likely to claim they had seen less important actions than more important actions. In a later study, Gerrie et al. showed another group of participants the sandwich movie with either all the most crucial or all the least crucial actions missing. At a recognition test, participants were more likely to think they had seen the missing least crucial information than the missing most crucial information.

Gerrie et al. (2006) explained their results using the source-monitoring framework (SMF; Johnson et al., 1993). According to the SMF, people use the qualitative characteristics of some recalled information to determine its source, to decide whether it is real or imagined, true or false, or—as in the Gerrie et al. case—old or new. They hypothesised that if missing clips were processed fluently at test, they would feel like an old clip. That is, participants would experience a sudden rush of familiarity, and—in the absence recalling that a given clip was missing earlier—would often misjudge it as an old one (Jacoby, Kelley & Dywan, 1989; Whittlesea, Masson, & Hughes, 2005; Whittlesea & Williams, 2000). However, if participants noticed a particular clip was missing at study, then at test they would rely on their recollection that they did not see that clip. Thus, participants’ source memory could override the effects of a feeling of familiarity (Rotello & Heit, 2000).

It stands to reason, then, that factors affecting people’s source-monitoring ability can affect the way they remember the movie clips. One factor that should affect source-monitoring ability is working memory capacity (WMC). An accumulating body of research suggests that WMC is strongly related to one’s ability to control attention (Conway, Cowan, & Bunting, 2001; Engle, 2002; Engle & Kane, 2004; Hester & Garavan, 2005; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2000, 2003). Put another way, WMC is related to people’s ability to marshal cognitive resources to track important information and shut out distracting information. Thus far WMC has been linked to false memories in two paradigms: the misinformation effect and the DRM effect.

In a typical three-stage misinformation study, participants watch an event, then read some erroneous information about the event, and then report what they remember seeing. Participants are said to be misled when they report having seen information in the event that was actually part of the postevent narrative (Loftus, Miller, & Burns, 1978; Loftus, 2005). Various factors relate to people’s susceptibility to the misinformation effect, including age (Bruck & Ceci, 1999; Karpel, Hoyer, & Toglia, 2001), dissociative tendencies (Eisen & Lynn, 2001), and WMC (Jaschinski & Wentura, 2002).

Jaschinski and Wentura (2002) measured people’s WMC using operation span (OSPAN; Turner & Engle, 1989). In an OSPAN task, participants must remember letter strings while calculating mathematical operations (e.g., \(9 - 4 \times 2 = 10\); True or False?). Higher OSPAN scores indicate a higher WMC (known as high spans), while lower OSPAN scores indicate a lower WMC (known as low spans). Jaschinski and Wentura found a negative correlation \(r = -0.39\) between participants’ OSPAN scores and the extent to which they were misled by the postevent information.

More recently, Watson, Bunting, Poole, and Conway (2005) found that WMC was related to participants’ ability to heed a warning about the DRM effect and resist false memories at test. When participants were not warned, there was no difference between high and low spans on either correct recall of old words or false recall of lure words.
Taken together, these studies suggest that people’s WMC is related to their source-monitoring ability. In this paper we examine whether high spans are better able to ward off false memories for missing information than low spans. Our rationale was straightforward. Successful performance on our task demands that participants first watch the event and notice that some actions are missing. Moreover, noticing whether a specific action was missing depends on its crucialness; perhaps unsurprisingly, people tend to notice when crucial information is missing more often than when noncrucial information is missing (Gerrie, 2006; Gerrie et al., 2006). At test, the actions that belong to the event—both the seen and missing actions—should feel familiar, while control clips should not. But familiarity is sometimes an untrustworthy guide, offering no helpful information about whether or not participants have seen the action (see Jacoby et al., 1989). Thus, to do well on the test means that participants must first notice that a clip was missing from the event, and second overcome the undependable feeling of familiarity at test.

These processes lead us to expect two general patterns of results. First, we should find that participants falsely recognise fewer crucial clips (see also Gerrie et al., 2006). Second, participants with higher WMC should be better able to keep track of which clips were missing while warding off the irrelevant feeling of familiarity at test. Thus, high spans should be less likely to falsely recognise missing crucial clips than low span participants. If higher spans falsely recognised missing crucial information less often than lower spans, it would be additional evidence that WMC is related to source-monitoring behaviour.

In Study 1a we adapted a method by Reed, Montgomery, Schwartz, Palmer, and Pittenger (1992) to determine which parts of four everyday events were crucial or not crucial. Then, in Study 1b, we used automated OSPAN task to screen a large group of participants for their WMC (Turner & Engle, 1989; Unsworth, Heitz, Shrock, & Engle, 2005). Later, we used a median split to divide participants into high and low spans, and showed a sample of them the everyday events with either the crucial information missing, or the noncrucial information missing. A short time later we tested participants’ recognition memory for parts of the events they had seen (old clips), parts of the movie that had been removed (missing clips), and movie clips that did not belong in the events (control clips).

**STUDY 1A**

In Gerrie et al. (2006), participants saw a single event. Our goal in Study 1a was to analyse a new set of events, and determine which aspects were most and least crucial. To achieve this goal, we adapted the method developed by Reed et al. (1992), and asked participants to watch the events, and rate which segments were most and least crucial.

**Method**

**Participants**

A total of 44 introductory psychology students participated as observers in return for course credit.

**Procedure**

We filmed a woman performing four everyday events to get ready for work: making a sandwich, making a cup of coffee, making the bed, and brushing her teeth. We asked six raters to watch the video and list the major steps in each event; we used the most commonly identified steps to break the movie down into 48 clips (see Table 1 for descriptions). We then edited the movie using Apple iMovie so that each major action was separated by black screen for 1 second.

Next, we used Reed et al.’s (1992; Gerrie et al., 2006) method to distinguish between crucial actions and noncrucial actions. We gave the 44 observers the following instructions:

Most everyday actions are made up of action units. For example, when people cook a meal (an everyday action), they must perform several action units, heat the oven, prepare the food, put the food in the oven, remove the food from the oven etc…I am interested in the action units of a few everyday events.

In all everyday events, some action units are crucial to the outcome of the whole event, while others are not crucial. During this experiment, I would like to know which action units you think are crucial to each of the events you are shown.
Your task:

1. First I will play a video of a woman named Jenny performing some everyday tasks. Please watch the video carefully.
2. Next, I will play the same video broken down into shorter movie clips. Each short clip is an action unit.
3. Your task is to rate each action unit as either crucial or not crucial to the outcome of the whole event. Then indicate how confident you are about your selection by using a 5-point scale.

The observers watched the four events twice; first to familiarise them with the movie; second to complete the rating task. During the rating task, observers were shown each event one clip at a time and were asked to complete two questions. First, they indicated if the clip was crucial or not crucial to the overall event. Second, they were asked to rate their confidence in their rating on a 5-point scale where 1 was labelled “not at all confident” and 5 was labelled “very confident”.

Results and discussion

We took participants’ “crucialness” classifications and combined them with their confidence ratings to produce a new score. Clips rated as crucial were weighted +1; clips rated noncrucial were weighted −1, and weight × confidence produced the new score. Thus, each clip fell on a scale from +5 (crucial and very confident) to −5 (not crucial and very confident). Clips with a mean score +3 or higher were classified as crucial. Clips with a mean score −3 or less were classified as noncrucial. Ultimately, we derived 13 crucial and 10 noncrucial clips from this analysis. The middle two columns of Table 1 show, for each action unit clip, participants’ ratings and classification.

We used these classifications to develop two versions of the movie. In one version, we removed the noncrucial clips and called it the crucial present version; in the other version, we removed the crucial clips, and called it the crucial absent version. In other words, people who watched the crucial present version would never see noncrucial clips (e.g., cutting the sandwich in half; putting the lid back on the milk). Likewise, people who watched the crucial absent version would never see crucial clips (e.g., placing the two pieces of bread together; pouring the coffee). The right-hand columns in Table 1 show the role each clip played in each version.

Note that we made several adjustments when creating the movies. First, we did not remove a clip if it fell immediately before or after another with the same classification. For example, ratings from 1a showed that two sequential actions—placing the bedspread on the bed and the pillows on the bed—were both crucial (see Table 1). In such a case, we removed only one of those clips. Second, we did not remove a crucial or noncrucial clip if it was the first or last clip in an event. Finally, because none of the clips in the bed-making event was rated less than −3, we classified the clip with the lowest rating (smoothening the sheet, M = −2.7) as noncrucial, ensuring that we removed an equal number of crucial and noncrucial clips from each event. Ultimately, each version of the movie consisted of 48 clips, of which 9 were removed.

STUDY 1B

In Study 1b we used the event versions developed in Study 1a to examine the relationship between false memories for missing slides and WMC.

Method

Participants

A total of 76 introductory psychology students participated in all phases of the study in return for course credit.

Design

We used a 2 (OSPAN: high vs low span) × 2 (video version: crucial absent vs. crucial present) × 3 (clip type: old vs missing vs control) mixed design.

Procedure

Screening phase. We measured WMC among 899 participants during a departmental mass-testing session by using the automated OSPAN task (Sobel, Gerrie, & Garry, 2005; Unsworth et al., 2005) on Dell P4 2.8 GHz computers with 15-inch LCD monitors. In the OSPAN, participants see a mathematical operation stated as a proposition to be judged true or false. For example, participants might see $2 \times (2 + 6) = 15$,
### TABLE 1

Descriptions of clips, their ratings and classifications as determined in Study 1a

<table>
<thead>
<tr>
<th>Clip</th>
<th>Description</th>
<th>Mean rating in Study 1a</th>
<th>Classification</th>
<th>Missing clips in Crucial absent version</th>
<th>Missing clips in Crucial present version</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Take bread out of bag</td>
<td>4.60 (0.50)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Take lid off peanut butter</td>
<td>2.70 (3.11)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Spread peanut butter on bread</td>
<td>3.20 (2.93)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Put lid on peanut butter</td>
<td>−4.45 (0.76)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Take lid off jam</td>
<td>1.65 (3.69)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Spread jam on bread</td>
<td>3.00 (2.96)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Put lid on jam</td>
<td>−4.40 (0.60)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Place bread together</td>
<td>4.00 (1.81)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Cut sandwich in half</td>
<td>−3.75 (1.97)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Wipe knife with napkin</td>
<td>−3.85 (3.05)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Turn jug on</td>
<td>4.10 (2.20)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Take lid off coffee</td>
<td>2.80 (3.14)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Spoon coffee into cup</td>
<td>4.70 (0.57)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Put lid on coffee</td>
<td>−3.85 (2.18)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Pour water into plunger</td>
<td>4.20 (2.21)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Put lid on plunger</td>
<td>−0.15 (3.86)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Lid off sugar</td>
<td>−0.60 (3.89)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Spoon sugar in cup</td>
<td>−0.10 (3.99)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Put lid on sugar</td>
<td>−4.05 (2.21)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Take lid off milk</td>
<td>−0.45 (4.03)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Pour milk in cup</td>
<td>0.40 (4.01)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Put lid on milk</td>
<td>−3.80 (2.21)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Plunge coffee</td>
<td>3.30 (3.08)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Pour coffee in cup</td>
<td>4.70 (0.57)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Stir coffee</td>
<td>−1.05 (3.58)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Drink coffee</td>
<td>−1.75 (3.73)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Take blanket off bed</td>
<td>0.85 (3.54)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Take duvet off bed</td>
<td>−0.85 (3.84)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Take pillows off bed</td>
<td>−1.95 (3.39)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Smooth sheet</td>
<td>−2.70 (3.18)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Place duvet on bed</td>
<td>4.10 (1.80)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Place pillows on bed</td>
<td>3.95 (1.82)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Fold blanket</td>
<td>−1.00 (3.84)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Place blanket on bed</td>
<td>−1.40 (3.53)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Take toothbrush out</td>
<td>2.20 (3.56)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Open toothpaste</td>
<td>2.05 (3.73)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Squeeze toothpaste onto brush</td>
<td>3.55 (2.24)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Put toothpaste away</td>
<td>−3.60 (1.73)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Turn tap on</td>
<td>0.65 (3.48)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Hold toothbrush under tap</td>
<td>−1.30 (3.34)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Brush teeth</td>
<td>4.65 (0.75)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Spit out toothpaste</td>
<td>0.85 (3.65)</td>
<td>Crucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Rinse toothbrush</td>
<td>−1.40 (3.75)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Put toothbrush away</td>
<td>−3.05 (2.84)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Rinse mouth with water</td>
<td>−1.65 (3.79)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Spit out water</td>
<td>−0.45 (4.03)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Turn tap off</td>
<td>−0.65 (3.98)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Wipe mouth with towel</td>
<td>−3.35 (2.32)</td>
<td>Noncrucial</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Numbers are mean scores with standard deviations in brackets.
and should use a mouse button to click “false”. Next, they see a letter on the screen (such as “K”) then another operation, and so on until the end of each trial. At the end of each trial, participants see a grid of 12 letters, and use the mouse to indicate, in order, which letters they saw. Each participant has a response deadline, based on performance in a set of practice trials at the beginning of the task.

In our adaptation, participants completed 75 trials consisting of three to seven letters. They were instructed to keep their operation performance above 85% and were warned by the computer when they dropped below that criterion at any point. We calculated OSPAN scores as the sum of all letters correctly recalled from correct trials. We excluded participants whose overall performance was below 85%, and participants who did not complete all 75 trials.

OSPAN scores were comparable to other published studies using the automated OSPAN, ranging from 0 to 75 with a median score of 40 (see Unsworth et al., 2005). All participants who scored 40 or higher were classified as high spans and all participants who scored below 40 were classified as low spans.

Between 2 and 8 weeks later, a new experiment was made available to all mass-testing participants. Participants who returned to complete the rest of the experiment were told that we were studying how people perceive and comprehend everyday events. There were two phases, a study phase and a test phase.

**Study phase.** In the study phase, participants were randomly assigned to watch either the crucial present or the crucial absent version of the movie. Next, participants completed word puzzles for 20 minutes. The movies were played on a 25-inch television screen from a DVD.

**Test phase.** After the 20-minute filler task, all participants completed the same memory test. First, participants were given these instructions:

Earlier, you saw a woman named Jenny performing some tasks. Now you will be shown some short movie clips of Jenny performing some tasks, some of which you have seen before and others that you have not. Your task is to distinguish between the movie clips you have seen and those that you have not.

You will be shown each movie clip and then be asked to answer two questions. Firstly, you will be asked if you have seen each movie clip. You can respond by circling “Yes” or “No”.

Secondly, you will be asked how confident you are that you saw or did not see each movie clip. You may respond by circling 1, 2, 3, 4, or 5, where 1 means you are not at all confident, and 5 means you are very confident.

The recognition test consisted of 27 clips. Of those 27 clips, 9 were taken from the movie seen in the study phase (old clips), and 9 were the clips removed from the movie in the study phase (missing clips). Old clips from the crucial present version served as missing clips in the crucial absent version, and vice versa. Nine control clips showed the same woman from the movie in the study phase performing actions that could plausibly have happened, but broke the structure of the events seen earlier. For example, control clips showed the woman pouring a juice while making the sandwich or wearing a different-coloured top while brushing her teeth.

**Results and discussion**

Recall that the primary aim of Study 1b was to examine the relationship between WMC and false memories for missing slides. We present our results in two parts. First, we examine participants’ correct and false recognition as a function of both operation span and the presence (or absence) of crucial information. Next, we examine participants’ confidence for their memories.

**Recognition performance**

Was WMC associated with recognition of crucial and noncrucial information? To address this question we first calculated, for each participant, the proportion of “yes” responses to old, missing, and control clips. We classified these responses according to whether participants were high or low span, and furthermore whether they saw the “crucial present” or the “crucial absent” version of the movie. These data appear in Figure 1.

Figure 1 shows three important ways that recognition depended on WMC. First, the black bars on both sides of the figure show that people were equally accurate at recognising old clips, regardless of their WMC, or the importance of the clips. Second, the grey bars on the left side of the figure show that low spans were more likely to falsely recognise crucial information than high...
spans. Third, the grey bars on the right side of the figure show that WMC was unrelated to false recognition of noncrucial information.

Put another way, we found a three-way interaction among clip type, OSPAN, and video version, $F(1, 72) = 4.86, p = .03$; Partial $\eta^2 = .06$. Moreover, follow up $t$-tests showed that there were no differences between high and low spans’ correct recognition of crucial information ($M = .85, SD = .13$ and $M = .86, SD = .14$ respectively) $t(36) = .13, ns$, or noncrucial information ($M = .89, SD = .09$ and $M = .90, SD = .12$ respectively) $t(36) = .17, ns$. In addition, low spans were more likely to falsely recognise missing crucial clips ($M = .29, SD = .25$) than high spans ($M = .11, SD = .10$; $t(36) = 2.85, p < .01$; Cohen’s $d = .95$); low spans falsely recognised missing noncrucial clips ($M = .55, SD = .18$) at the same rate as high spans ($M = .60, SD = .26$) $t(36) = .64, ns$.

Recently, Conway et al. (2005) urged caution in using statistical analyses that rely on extreme OSPAN scores, such as our median split, and recommended capitalising on the continuous range of scores where possible (see also MacCallum, Zhang, Preacher, & Rucker 2002). In line with this recommendation, we used the OSPAN as a continuous variable and conducted a moderation analysis (a form of multiple regression; see Baron & Kenny, 1986). The interaction between clip importance and OSPAN significantly predicted false recognition ($Beta = .76, p = .02$; $\Delta R^2 = .04$ or $adj R^2 = .46$), a finding that replicated the pattern of results we reported using an ANOVA on the median split, and—for ease of interpretation—we concentrate on the ANOVA results here.

In addition to performing the moderation analysis, we also examined the correlation between OSPAN and false recognition of both crucial and noncrucial clips. We calculated, for each participant, the proportion of “yes” responses to missing clips and plotted these responses against participants’ OSPAN scores, according to whether they saw the “crucial present” or the “crucial absent” version of the movie. These data appear in Figure 2. There were two important patterns. First, the left panel shows that people’s WMC was related to how often they falsely recognised crucial information. Second, the right panel of the figure shows that their WMC was not related to how often they falsely recognised noncrucial clips. In other words, there was a significant negative correlation between OSPAN and false recognition of missing clips in the CA condition ($r = -.40, p = .01$), but not in the CP condition ($r = .15, ns$).

![Figure 1](image1.png) Mean proportion of “Yes” responses for each clip type as a function of video version and OSPAN in Study 1b.

![Figure 2](image2.png) Scatterplots showing the mean proportion of false recognition plotted against OSPAN score in the Crucial absent and Crucial present conditions in Study 1b.
In sum, these data suggest that WMC is related to people’s ability to monitor the source of memories for events. However, what we do not know is whether WMC is related to different subjective experience of these memories. To answer this question, we examined participants’ confidence judgements in two ways. First we examined their overall confidence, regardless of whether they were correct or incorrect. Second we examined their confidence for “Yes” responses only; that is, we compared confidence for true memories with confidence for false memories.

**Overall confidence**

We calculated a mean confidence rating for each participant’s response to old and missing clips. We classified these responses according to whether participants were high or low span, and furthermore whether they saw the crucial present or the crucial absent version of the movie. These data appear in Figure 3.

Figure 3 shows three findings about people’s overall confidence. First, high spans were more confident of their responses than low spans, regardless of the importance of the clips. In other words, there was a main effect for OSPAN, $F(1, 73) = 5.28, p = .02; \text{Partial } \eta^2 = .07,$ in that high spans ($M = 4.29; SD = .57$) were more confident of all their responses than low spans ($M = 4.06; SD = .53$). Participants’ WMC was not significantly related to any other variables.

Second, the right side of the figure shows that people were more confident of old clips than missing clips for the crucial present version of the movie. Finally, the left side of the figure shows people were equally confident for old and missing clips for the crucial absent version of the movie. In other words, we found a two-way interaction between clip type and video version, $F(1, 73) = 43.46, p < .01; \text{Partial } \eta^2 = .37.$

**Confidence for true vs false recognition**

Next we examined people’s confidence for “Yes” responses: confidence of true versus false recognition. We calculated a mean confidence rating for each participant’s “Yes” response to old and missing clips. We classified these responses according to whether participants were high or low span, and furthermore whether they saw the “crucial present” or the “crucial absent” version of the movie. These data appear in Figure 4.

Figure 4 shows two findings about people’s confidence for true and false recognition. First, people’s WMC did not affect their confidence for either true or false memories. In other words, WMC did not significantly affect any other variable when examining confidence for “Yes” responses. Second, people were more confident of their true memories than of their false memories regardless whether the information was crucial or noncrucial. In other words, people were more confident of a “Yes” response to old clips ($M = 4.48, SD = .39$) than to missing clips ($M = 3.62,$

![Figure 3](image1)

**Figure 3.** Mean overall confidence rating for each clip type as a function of video version and OSPAN in Study 1b.

![Figure 4](image2)

**Figure 4.** Mean confidence rating for “Yes” responses for each clip type as a function of video version and OSPAN in Study 1b.
\[ SD = .80 \] as shown by a main effect for clip type, \( F(1, 65) = 97.66, p < .01; \) \( \text{Partial } \eta^2 = .60. \)

GENERAL DISCUSSION

The primary purpose of these studies was to ask whether WMC would be related to susceptibility to false memories for missing event information. The answer is a qualified yes: we found that high spans were better able to ward off false memories for crucial information, but not for noncrucial information. Additionally, we found that WMC had little effect on participants’ subjective experience of their memories, regardless of whether the information was crucial or not crucial. In other words, people’s WMC did not affect the confidence they had in their memories.

Our results fits with those found by Jaschinski and Wentura (2002), as well as Watson et al. (2005), who found that higher WMC reduced false memories in the misinformation effect and the DRM respectively. Our results also extend their findings by showing that higher WMC can reduce false memories for some types of information, but not others. Taken together, their results and ours suggest that people with higher WMC are better able to marshal resources that keep track of source-specifying information, especially in the face of distraction. For example, in a misinformation effect study, participants must distinguish between what they saw and erroneous information they read later. In Watson et al.’s version of the DRM, participants (particularly the warned ones) had to remember the words from the list, while trying to identify—and then ward off—the distracting influence of the lure. In our study, participants had to remember the clips from the movie, while ignoring the distraction that came from noticing that certain clips were missing. In all of these tasks, success requires people to keep their “eye on the prize”, to remember information accurately while at the same time monitoring sources of distraction.

There are at least two alternative explanations for the pattern of results we report here. First, high spans might be more motivated to perform more accurately than low spans. If this explanation were correct, high spans would not necessarily be better at source monitoring. Rather, they may simply have a greater desire to perform well in the task. Second, it is possible that high spans might simply be better at noticing missing crucial information. However, we do not find either of these explanations compelling. In either case, we would expect high spans to recognize old clips more often than low spans. However, that is not what we found. Instead, we found participants’ recognition of old clips did not differ, regardless of WMC. Thus, there is no evidence for either the motivation or the “notice at event” explanations.

It is important to note that we manipulated the “crucialness” of information between participants (as did Gerrie et al., 2006). That is, participants saw a movie in which either all the most crucial parts or all the least crucial parts were removed. Another way of describing our design is to say that the “crucial present” vs “crucial absent” manipulation was perfectly confounded with a “non-crucial absent” vs “non-crucial present” manipulation. As such, it is possible that we have unwittingly compared events with different meanings—we cannot determine whether our effects are driven by what was participants saw or what they did not see. The question is one worthy of future study. One fruitful approach would be to manipulate “crucialness” within participants, by removing half the crucial and half the noncrucial clips. If our effects were caused by differences in meaning, then we should find that false recognition for crucial and non-crucial clips was equal.

There are theoretical and practical implications for our results. Theoretically, our research contributes to the accumulating body of literature showing that characteristics of events are remembered differently, depending on the characteristics of the rememberer. Specifically, our results add to those showing that people with high WMC ward off distraction by controlling their attention (Conway et al., 2001; Engle, 2002; Engle & Kane, 2004; Hester & Garavan, 2005; Kane et al., 2001; Kane & Engle, 2000, 2003). More practically, our results suggest a means by which people draw on cognitive abilities to adjust their source monitoring. For example, Mitchell and Johnson (2000) propose that—depending on the task at hand—people change their criteria for the characteristics that must be present for them to decide that a memory comes from a real experience, rather than from their imagination. They noted that being an eyewitness in court would invoke stricter criteria than merely recalling who told a funny joke. Our results suggest that people’s ability to
adjust these criteria might depend on their working memory capacity.

References


Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented...


