Increased False-Memory Susceptibility After Mindfulness Meditation

Brent M. Wilson¹, Laura Mickes², Stephanie Stolarz-Fantino¹, Matthew Evrard¹, and Edmund Fantino¹
¹Department of Psychology, University of California, San Diego, and ²Department of Psychology, Royal Holloway, University of London

Abstract
The effect of mindfulness meditation on false-memory susceptibility was examined in three experiments. Because mindfulness meditation encourages judgment-free thoughts and feelings, we predicted that participants in the mindfulness condition would be especially likely to form false memories. In two experiments, participants were randomly assigned to either a mindfulness induction, in which they were instructed to focus attention on their breathing, or a mind-wandering induction, in which they were instructed to think about whatever came to mind. The overall number of words from the Deese-Roediger-McDermott paradigm that were correctly recalled did not differ between conditions. However, participants in the mindfulness condition were significantly more likely to report critical nonstudied items than participants in the control condition. In a third experiment, which tested recognition and used a reality-monitoring paradigm, participants had reduced reality-monitoring accuracy after completing the mindfulness induction. These results demonstrate a potential unintended consequence of mindfulness meditation in which memories become less reliable.

Keywords
false memories, mindfulness, Deese-Roediger-McDermott (DRM) paradigm, source monitoring, reality monitoring, signal detection theory

Received 10/22/14; Revision accepted 6/9/15

The concept of mindfulness is pervasive in both popular culture and academic research. Oprah Winfrey, Deepak Chopra, and Dr. Oz (The Dr. Oz Show, 2013) have all extolled the merits of being mindful, and scholarly studies have investigated the benefits of this phenomenon. Mindfulness-based interventions for both physical and psychological disorders have been reported, and these include reduced pain intensity for patients with chronic pain (Reiner, Tibi, & Lipsitz, 2013), improved psychological well-being (Brown & Ryan, 2003), reduced levels of stress and anxiety (Astin, 1997; Jain et al., 2007; Rosenzweig, Reibel, Greeson, Brainard, & Hojjat, 2003; Shapiro, Schwartz, & Bonner, 1998), and decreased depression in older adults (Geschwind, Peetters, Drukker, van Os, & Wichers, 2011). Mindfulness meditation focuses attention on the present moment in an accepting and nonjudgmental manner (Baer, Smith, & Allen, 2004; Brown & Ryan, 2003; Kabat-Zinn, 2013). Each thought, feeling, and sensation is acknowledged and accepted without judgment or evaluation (Bishop et al., 2004; Kabat-Zinn, 2013; Segal, Williams, & Teasdale, 2012; Teasdale, 1999). As Kabat-Zinn (2013) noted, “the practice involves suspending judgment and just watching whatever [emphasis in original] comes up” (p. 23).

In contrast to the myriad benefits of mindfulness, it may also increase false-memory susceptibility by affecting the cognitive operations needed to distinguish between internal and external sources of information. According to the source-monitoring framework, false memories occur because of a failure to distinguish the origin of a memory (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 2008). When the origin of a memory is misattributed, information from one context is falsely remembered as having

Corresponding Author:
Brent M. Wilson, Department of Psychology, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093
E-mail: bwilson@ucsd.edu
been part of a different context. Source-monitoring errors can arise as a consequence of confusing memory sources. Confusion can occur between two external sources as well as between an internally generated source and an external one (Johnson et al., 1993).

Reality monitoring is the process of discriminating between internally generated and external memory sources (Johnson & Raye, 1981). Information that people generate themselves is usually associated with cognitive operations (i.e., mental processes involved in the generation of information) that leave a trace and later provide cues that the information was internally generated rather than actually encountered in the external world (Johnson, Raye, Foley, & Foley, 1981; Lindsay, 2008). If focusing mindful attention without judgment results in the suspension of cognitive operations (and thus the elimination of the trace records those operations would otherwise leave), people will have greater difficulty differentiating internal and external sources of information. That is, mindfulness training might increase the risk for false memories because internally generated memories would lack the cues that are ordinarily used to help identify them as having been internally generated.

In the first two experiments, we examined the effect of mindfulness meditation on false-memory susceptibility using the Deese-Roediger-McDermott (DRM) paradigm (Roediger & McDermott, 1995). The DRM is the paradigm most widely used to test false memories (Brainerd & Reyna, 2005). The procedure involves presenting lists of closely related words and then testing memory with either recall or recognition. For each list, there is a word (the critical item) that is closely related to the words on the list but is not on the list. The critical item is strongly activated by the other words on the list, and it can be falsely remembered if people mistake this strong internal activation for an actual memory of the word. For example, the word list garbage, waste, can, refuse, sewage, bag, junk, rubbish, sweep, scraps, pile, dump, landfill, debris, and litter can activate the critical item trash (list from Roediger, Watson, McDermott, & Gallo, 2001).

In the third experiment, we used a reality-monitoring paradigm and extended the research to recognition memory. If increases in false memories after mindfulness training are due to reduced reality-monitoring abilities, participants will have reduced abilities to discriminate between words actually studied and words internally activated during study but not actually presented.

## Experiment 1

### Method

**Participants.** One hundred fifty-three undergraduate students (37 male, 116 female; mean age = 20.7 years, \(SD = 2.4\)) at the University of California, San Diego, participated in this experiment for course credit. We planned to recruit as many participants as possible before the end of the quarter.

**Materials and procedure.** Participants sat in individual sound-attenuated rooms and were randomly assigned to receive either a 15-min mindfulness induction or a 15-min mind-wandering induction. In the mindfulness induction, participants listened to a guided focused-breathing exercise recorded by Marilee Bresciani Ludvik at the Rushing to Yoga Foundation. This mindfulness induction was based on a script by Arch and Craske (2006) that had been adapted from work by Kabat-Zinn (1990). It instructed participants to focus attention on their breathing without judgment. The mind-wandering induction, also recorded by Marilee Bresciani Ludvik, instructed participants to think about whatever came to mind. Mind wandering has been used as a control condition in other mindfulness experiments to represent a neutral mental state (e.g., Hafenbrack, Kinias, & Barsade, 2014; Kiken & Shook, 2011).

All participants were then shown the DRM word list for the critical item trash (Roediger et al., 2001). Each word was presented in the center of the computer screen for 1.5 s. After all 15 words were presented, participants immediately typed as many words as they could remember.

### Results

Participants in the mindfulness condition were significantly more likely to falsely remember seeing the word trash, 39%, 95% confidence interval, or CI = [29.15%, 49.46%], than those in the mind-wandering condition, 20%, 95% CI = [12.37%, 31.35%], \(z = 2.48, p = .014\), Cohen’s \(d = 0.50, 95\% \text{ CI} = [0.18, 0.82]\). The mean number of correctly recalled words did not significantly differ between the mindfulness condition, 7.02, 95% CI = [6.68, 7.37], and the mind-wandering condition, 6.75, 95% CI = [6.35, 7.15], \(t(152) = 1.02, p > .250\). For each list, we numbered the recalled words according to the order in which they were recalled. The average position number at which the critical item was reported did not differ significantly between the mindfulness condition (6.3) and the mind-wandering condition (6.1), \(t(45) = 0.2, p > .250\). The average number of other words falsely recalled did not significantly differ between the mindfulness condition (0.34) and the mind-wandering condition (0.29), \(t(152) = 0.45, p > .250\).

## Experiment 2

### Method

**Participants.** One hundred forty undergraduate students (40 male, 100 female; mean age = 21.5 years, \(SD = 4.3\)) at the University of California, San Diego,
participants in this experiment for course credit. Using our effect size from Experiment 1, we estimated that we would need 128 participants to have 80% power to detect a statistically significant difference. We planned to recruit as many participants as possible before the end of the quarter, but the minimum was 128 participants.

Materials and procedure. Participants sat in individual sound-attenuated rooms. Six (preinduction) DRM word lists (critical items: mountain, music, thief, doctor, cold, needle) from Roediger et al. (2001) were presented in random order. Each word was presented in the center of the computer screen for 1.5 s. After viewing each list, participants immediately typed as many words as they could remember.

After the six lists were completed, the computer randomly assigned participants to either the mindfulness condition or the mind-wandering condition. The inductions were those used in Experiment 1. Participants then completed a different set of six postinduction DRM word lists (critical items: lamp, trash, slow, wish, foot, window) also from Roediger et al. (2001) presented in random order. Each word was presented in the center of the computer screen for 1.5 s. Again, after viewing each list, participants immediately typed as many words as they could remember.

Results

In the within-subjects comparison, participants in the mindfulness condition were significantly more likely to falsely recall the critical items after the induction than before the induction, \( t(67) = 2.75, p = .008 \), Cohen’s \( d = 0.33 \), 95% CI = [0.09, 0.58]. Participants in the mind-wandering condition showed no difference in critical-item recall on the preinduction and postinduction lists, \( t(71) < 0.001, p > .250 \), Cohen’s \( d = 0.00 \), 95% CI = [0, 0]. The same results were also found in the between-subjects comparison. Participants in the mindfulness condition were significantly more likely to falsely recall the critical item, \( M = .34, 95\% \text{ CI} = [.29, .38] \), than were participants in the mind-wandering condition, \( M = .26, 95\% \text{ CI} = [.21, .31] \), \( t(138) = 2.27, p = .025 \), Cohen’s \( d = 0.38, 95\% \text{ CI} = [.05, 0.72] \). This difference remained significant when we controlled for participants’ baseline levels of false-memory susceptibility and memory performance using the average critical-item recall and proportion correct during preinduction, \( F(1, 136) = 5.78, p = .018 \). We performed a 2 × 2 analysis of variance and found a significant interaction between condition (mindfulness vs. mind-wandering) and time of recall (preinduction vs. postinduction), \( F(1, 138) = 4.22, p = .025 \), Cohen’s \( d = 0.38, 95\% \text{ CI} = [.05, 0.72] \). Figure 1 shows the average proportion of critical items falsely recalled as being included on the preinduction and postinduction word lists.

The average proportion of words correctly recalled did not differ significantly between conditions (mindfulness: \( M = .46, 95\% \text{ CI} = [.44, .49] \); mind-wandering: \( M = .45, 95\% \text{ CI} = [.43, .48] \), \( t(138) = 0.66, p > .250 \), Cohen’s \( d = 0.11, 95\% \text{ CI} = [-0.22, 0.44] \). The proportion correct was not significantly different even after we controlled for both correct identifications and critical-item recall on the preinduction lists, \( F(1, 136) = 1.66, p = .200 \). Participants in the two conditions did not significantly differ in critical-item recall (\( p > .250 \)) or correct recall (\( p > .250 \)) on the DRM lists completed before receiving the audio inductions. Again, for each list, we numbered the recalled words according to the order in which they were recalled. The average position number at which the critical item was reported did not differ significantly between the mindfulness condition (5.7) and the mind-wandering condition.
condition (5.2), k(119) = 1.60, p = .111, and did not change significantly after participants completed the mindfulness induction (5.4 for preinduction and 5.7 for postinduction), k(50) = 0.55, p > .250. The average number of other words falsely recalled did not differ significantly between the mindfulness condition (0.22) and the mind-wandering condition (0.18), k(138) = 0.96, p > .250, and did not change after participants completed the mindfulness induction (0.22 for both preinduction and postinduction), k(67) = 0, p > .250.

Discussion

These results provide evidence that false-memory susceptibility increases after completing mindfulness training. The pretest-posttest design of this experiment (as opposed to the design of Experiment 1) also provides evidence that false-memory susceptibility is increased by mindfulness training rather than being decreased by mind wandering. In the next experiment, we extend this work to a reality-monitoring paradigm (Brainerd & Reyna, 2005) to better identify why false memories increase after mindfulness meditation training.

Experiment 3

Method

Participants. Two hundred fifteen undergraduate students (59 male, 156 female; mean age = 20.3 years, SD = 2.9) at the University of California, San Diego, participated in this experiment for course credit. On the basis of the effect size from our within-subjects comparison in Experiment 2, we estimated that we would need 75 participants to have 80% power to detect a statistically significant difference. We planned to recruit as many participants as possible before the end of the quarter, but the minimum was 75 participants.

Materials and procedure. Two hundred pairs of strongly associated words (e.g., foot-shoe, sediment-fossil) were constructed using databases of word associations (Palermo & Jenkins, 1964; Rotmistrov, 2014). One hundred word pairs were randomly selected for the preinduction study and test phase. The remaining 100 word pairs were then used for the postinduction study and test phase.

Participants sat in individual sound-attenuated rooms. During the preinduction study phase, 1 word from each pair was randomly selected and presented in the center of the computer screen for 1.5 s. The 100 words were presented in random order. After all the words had been presented to participants, the preinduction test phase began immediately. One word from each pair was randomly selected for the test phase and presented in the center of the computer screen. This procedure gave each word an equal probability of being a target or a lure. Participants identified whether the word had appeared on the word list (“old”) or had not appeared on the word list (“new”) and indicated their level of confidence in each answer.

All participants then listened to the 15-min mindfulness induction used in the first two experiments. After completing the mindfulness induction, participants began the postinduction study phase followed immediately by the postinduction test phase. The procedure was identical to that in the preinduction study and test phase.

Results

We used d' (Macmillan & Creelman, 2005) to compare how well participants were able to discriminate between externally presented (old or target) items and internally generated (new or lure) items. Accuracy (d') was significantly higher for the word lists studied and tested before the mindfulness induction ($M = 1.60$, $SD = 0.71$) than for the word lists studied and tested after the mindfulness induction ($M = 1.42$, $SD = 0.79$), k(214) = 4.08, $p < .001$, Cohen's $d = 0.28$, 95% CI = [0.14, 0.41]. With regard to the proportion of words declared to be “old,” there was a significant interaction between the status of the word (internal vs. external) and condition (control vs. mindfulness), F(1, 214) = 20.94, $p < .001$. The false alarm rate increased significantly after participants completed the mindfulness induction (before: $M = .20$, $SD = .15$; after: $M = .25$, $SD = .18$), k(214) = 4.49, $p < .001$, Cohen's $d = 0.31$, 95% CI = [0.17, 0.44], but the hit rate did not change significantly (before: $M = .72$, $SD = .15$; after: $M = .71$, $SD = .16$), k(214) = 1.55, $p = .123$, Cohen's $d = 0.11$, 95% CI = [−0.03, 0.24]. Because null-hypothesis significance testing cannot provide evidence in favor of the null, we also calculated the Jeffreys-Zellner-Siow (JZS) Bayes factor for the nonsignificant change in the hit rate (Rouder, Speckman, Sun, Morey, & Iverson, 2009). This method gave 5.65:1 odds in favor of the null hypothesis.

We used c (Macmillan & Creelman, 2005) to measure response bias. Participants had a significantly more liberal response bias (i.e., more of the distribution exceeded the criterion line) after completing the mindfulness induction (before: $M = 0.15$, $SD = 0.40$; after: $M = 0.085$, $SD = 0.47$), k(214) = 2.61, $p = .0097$, Cohen's $d = 0.18$, 95% CI = [0.04, 0.31]. However, it is important to note that a change in measured bias does not necessarily entail a change in participants’ decision strategy (Wixted & Stretch, 2000).

Discussion

The results of Experiment 3 are consistent with the results from Experiments 1 and 2 and provide additional evidence that mindfulness training increased false-memory susceptibility. Experiment 3 also extends the findings to
recognition memory and to a reality-monitoring paradigm. These findings support the idea that the increase in false memories is due to a reduction in reality-monitoring accuracy. Each word on the study list strongly activates its paired word. Participants are less accurate at discriminating between associated words (internally generated) and words actually studied (external memory source) after completing the mindfulness induction.

**General Discussion**

Our research adds to and connects the literature on mindfulness meditation and false memories. Whereas the preponderance of research on mindfulness has focused on the beneficial aspects of this phenomenon (Chiesa, Calati, & Serretti, 2011), our study examines a potential adverse effect. When meditators embrace judgment-free awareness and acceptance, their reality-monitoring accuracy may be impaired, increasing their susceptibility to false memories.

Information encountered in the external world is expected to leave a trace record that contains greater sensory detail than information that is internally generated, and this difference in sensory content is one factor that facilitates the discrimination between internally and externally generated information. Johnson et al. (1981) also noted the importance of a second factor: cognitive operations associated with the internal generation of information at the time of encoding. At retrieval, a trace record of those cognitive operations ordinarily helps to identify internally generated information as having been internally generated. However, the nonjudgmental aspect of mindfulness meditation may be expected to reduce this important cue. The essential idea of mindfulness meditation is to observe without judgment or reaction (rather than performing cognitive operations on) whatever comes to mind. The elimination of cognitive operations would therefore have the effect of also eliminating a trace record of such operations that might otherwise help to discriminate between internally and externally generated information on a later memory test. The result would be a decreased ability to discriminate between sources of information (Johnson & Raye, 1981), thereby increasing susceptibility to the DRM false-memory effect.

This argument can be illustrated using a simple signal detection model of a task in which the participant’s goal is to discriminate between internally generated (new) and externally presented (old) information (as in Experiment 3). The $x$-axis in the model shown in Figure 2 ranges from strong evidence that a test item was internally generated (at the far left) to strong evidence that the test item was externally presented (at the far right). The distribution of evidence values for internally generated items in the control condition falls farther to the left than the distribution of evidence values for externally presented items. The difference between the two distributions is $d'$. As noted earlier, this discrimination is facilitated both by the sensory content of the memory trace (more detailed for externally presented items than for internally generated items) and by the record of cognitive operations associated with the generation of internally generated items. Thus, for example, a test item that falls to the far left (a strong evidence trace for internal generation) might be associated with limited sensory content as well as a trace record of cognitive operations associated with the internal generation of that item. However, in the mindfulness condition, the trace record of cognitive operations is largely reduced. This reduction shifts the distribution associated with internally generated items to the right and increases the false alarm rate (i.e., the proportion of the internal distribution that falls above the decision criterion). A test item that falls to the far right, by contrast, might be associated with considerable sensory content and would also have no trace record of cognitive operations associated with internal generation (because the item was externally presented). Mindfulness, which selectively reduces cognitive operations, would therefore not change the representation of externally presented items, so the same external distribution would apply in both the control and the mindfulness conditions. If the decision criterion remains fixed across conditions, this
increase in the false alarm rate would not be accompanied by a change in the hit rate associated with externally generated items. Thus, the selective change in the false alarm rate would affect measured bias (more liberal in the mindfulness condition than in the control condition) even though the decision criterion remained unchanged.

Measured bias reflects the distance of the criterion line (i.e., the point at which participants switch from responding “new” to responding “old”) from the point of intersection for the internal and external distributions. In Figure 2, the point of intersection for the internal and external distributions in the mindfulness condition is farther to the left of the point of intersection for the internal and external distributions in the control condition. This means that the relative position of the criterion line stays the same in this model. Thus, the model predicts that measured bias should be more liberal for the mindfulness condition than for the control condition because of this change in the relative location of the decision criterion (resulting from an increase in the mean of the internal distribution in that condition). This simple model accounts for all of the results observed in Experiment 3, and it explains why false-memory susceptibility increases after mindfulness meditation.

A simple criterion-shift model (in which the distributions remain in the same locations but the criterion line changes) cannot fully account for the Experiment 3 results. Not only did measured bias change between conditions, $d'$ values also changed between conditions. The lower $d'$ value in the mindfulness condition means that the internal and external distributions moved in a manner that resulted in greater overlap between the two distributions. A simple criterion-shift model can explain only the change in measured bias; it cannot explain the change in $d'$ values observed between conditions.

Another possible model assumes that the effect occurs at retrieval rather than during encoding. Such a model can explain the change in $d'$ values but cannot readily explain all of the Experiment 3 results. According to this retrieval-based interpretation, one might assume that participants in the mindfulness condition respond on the basis of familiarity without engaging in recollection of source information (whereas control participants do engage in recollection of source information). In the absence of recollection, the internal distribution in the mindfulness condition would be to the right (in the external direction) of the internal distribution in the control condition because recollection would not count as evidence against familiar-but-imagined items having appeared on the list. By contrast, the external distribution in the mindfulness condition would be to the left (in the internal direction) of the external distribution in the control condition because recollection would not add evidence in favor of target items having appeared on the list. Thus, $d'$ values would be lower for the mindfulness condition, consistent with our results. However, the simplest version of this account would predict a difference in both hit and false alarm rates across conditions with no effect on measured bias, whereas we observed a selective effect on the false alarm rate and a clear effect on measured bias.

Mindfulness meditation appears to reduce reality-monitoring accuracy. By embracing judgment-free awareness and acceptance, meditators can have greater difficulty differentiating internal and external sources of information. As a result, the same aspects of mindfulness that create countless benefits can also have the unintended negative consequence of increasing false-memory susceptibility.

**Author Contributions**

The initial study concept came from B. M. Wilson and was developed by all authors. B. M. Wilson analyzed the data, and L. Mickes contributed to the data analysis. B. M. Wilson drafted the manuscript, and L. Mickes, S. Stolarz-Fantino, and E. Fantino edited it. All authors approved the final version of the manuscript for submission.

**Acknowledgments**

We thank John Wixted for valuable assistance with the signal detection model and data interpretation.

**Declaration of Conflicting Interests**

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

**Note**

1. These lists were not counterbalanced between preinduction and postinduction (which is not ideal for the within-subjects comparisons) because the preinduction lists were originally included to serve as covariates in the analysis of the postinduction word lists.

**References**


Increased False-Memory Susceptibility


