Exploring the Impact of Mindfulness on False-Memory Susceptibility

Susan M. Sherman and James A. Grange
School of Psychology, Keele University

Abstract

Wilson, Mickes, Stolarz-Fantino, Evrard, and Fantino (2015) presented data from three well-powered experiments suggesting that a brief mindfulness induction can increase false-memory susceptibility. However, we had concerns about some of the methodology, including whether mind wandering is the best control condition for brief mindfulness inductions. Here, we report the findings from a preregistered double-blind randomized controlled trial designed to replicate and extend Wilson et al.'s findings. Participants ($N = 287$) underwent 15-min mindfulness or mind-wandering inductions or completed a join-the-dots task before being presented with lists of words related to nonpresented critical lures. This was followed by free-recall and recognition tasks. There was no evidence for an effect of state of mind on correct or false recall or recognition. Furthermore, manipulation checks revealed that mindfulness and mind-wandering inductions activated overlapping states of mind. Exploratory analyses provided some support for mindfulness increasing false memory, but it appears that mind wandering may not be the right control for brief mindfulness research.

Keywords

false memory, mindfulness, mind wandering, replication, open data, open materials, preregistered

Wilson, Mickes, Stolarz-Fantino, Evrard, and Fantino (2015) reported that a brief mindfulness induction increased false-memory susceptibility. Media reports soon circulated of "how mindfulness plays havoc with memory" (Knapton, 2015). However, close reading of the research and of two subsequently published research articles—one supporting Wilson et al. and one contradicting them—suggests that this conclusion may be premature.

The Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995) is an effective method for eliciting false memories in the laboratory. Participants are presented with lists of words (e.g., *bed*, *rest*, *awake*) associated with a nonpresented critical lure (e.g., *sleep*). On subsequent memory tasks—such as free recall or recognition—participants typically have false memories for the nonpresented lure as well as true memories for presented list items.

Wilson et al. (2015) used the DRM paradigm to explore the effect of mindfulness and mind-wandering inductions on false recall. In Experiment 1, they used the relevant mind induction before showing participants words related to the nonpresented critical lure *trash* and giving them a free-recall task. Mindfulness induction led to 39% false recall, and mind-wandering induction led to 20% false recall. Wilson et al. suggested that mindfulness increases false-memory susceptibility. However, the authors did not include a baseline condition of no induction.

In Experiment 2, Wilson et al. (2015) explored the possibility that mindfulness increases false recall rather than that mind wandering reduces it. They presented participants with DRM lists before and after the mind inductions and compared pre- and postinduction recall performance. In the mind-wandering condition, false recall was the same before and after induction, but in the mindfulness condition, false recall increased after

Corresponding Author:
Susan M. Sherman, Keele University, School of Psychology, Dorothy Hodgkin Building, Keele, Staffordshire, ST5 5BG, United Kingdom
E-mail: s.m.sherman@keele.ac.uk

induction. However, there was again no baseline control condition, so it is difficult to know whether mindfulness and mind wandering increase or reduce false memories relative to no induction, especially given two other minor methodological concerns: (a) The pre- and postmanipulation lists were not counterbalanced, and (b) the backward associative strength (BAS) of the sets was not matched. BAS—defined as "the average tendency for words in the study list to elicit the critical item on a free association test" (Roediger, Watson, McDermott, & Gallo, 2001; p. 387)—is a key predictor of false recall. In Wilson et al.'s Experiment 2, the BAS of the preinduction lists was higher (range = 0.100–0.353, $M = 0.214$) than that of the postinduction set (range = 0.006–0.184, $M = 0.115$); so absent any mind induction, one would expect higher levels of false recall in the preinduction lists.

Using mind wandering as a control condition for mindfulness would be less problematic if we knew what the effect of mind wandering on false recall should be. To our knowledge, no previous work has addressed this. Mrazek, Franklin, Phillips, Baird, and Schooler (2013) define mind wandering as "a shift of attention from a task to unrelated concerns" (p. 776). Most theories of false memory posit that some encoding of the list items needs to take place for false memories to be facilitated—for example, via spreading activation (activation–monitoring account; e.g., Roediger et al., 2001) or gist extraction (fuzzy-trace theory; Reyna & Brainerd, 1995). According to these theories—and previous findings—the fewer items to be encoded, the fewer false memories created (e.g., Robinson & Roediger, 1997). If mind wandering is effectively induced so that participants are shifting their attention from the task (encoding the list items) to unrelated concerns, these theories might reasonably predict a drop in false memories rather than an increase. Alternatively, response bias—for example, when participants believe that they need to provide a certain number of answers but are unable to accurately recall or recognize enough of them—might result in increased false memories.

It is not clear that mind wandering was successfully induced in Wilson et al.'s (2015) study. If participants had shifted their attention away from the task, why was there no difference in their performance on correct recall of presented items between the mind-wandering and mindfulness conditions in either experiment? Risko, Anderson, Sarwal, Engelhardt and Kingstone (2012) provided evidence that correct memory should be impaired by mind wandering by observing that as mind wandering increased during a lecture, memory for the lecture material decreased.

A final methodological observation is that there was no measure of whether the mindfulness and mind-wandering inductions actually induced different mental states in participants. Several scales have been developed to measure either state or dispositional mindfulness, and although it would be nice to believe that an experimental induction works, one cannot assume that it does. Indeed, Brown and Ryan (2003) observed that both dispositional and state mindfulness vary across participants.

Since 2015, two further articles have been published exploring the impact of a brief mindfulness manipulation using the DRM paradigm. Experiment 2 by Rosenstreich (2016) used brief (30-min) mindfulness and mind-wandering manipulations and found that both correct and false recognitions increased in the mindfulness condition, supporting Wilson et al.'s (2015) findings. However, only 40 participants took part in this between-participants study, there was no baseline condition, and the effectiveness of the mindfulness manipulation was not measured. Baranski and Was (2017) explored the effects of 15-min mindfulness and mind-wandering manipulations and warning versus no-warning instructions on false memories. Following Wilson et al.'s method in Experiment 2, Baranski and Was had participants in their second experiment (a) study six DRM lists, each followed by free recall; (b) receive the induction; and (c) then study six more lists with free recall. They used three inductions: mindfulness, mind wandering, and puzzle completion. There was no difference in the amount of false recall between the conditions, and in all conditions, false recall declined after the manipulation. Exploratory analyses suggested that this decline was greatest in the mindfulness condition. These findings thus conflict with Wilson et al.'s; however, Baranski and Was also did not measure whether mindfulness was induced.

In light of the methodological issues identified in Wilson et al. (2015), we preregistered a study to replicate and extend Wilson et al.'s Experiment 1. Specifically, we (a) evaluated participant mindfulness before and after induction; (b) evaluated mind wandering after induction; (c) included mindfulness, mind-wandering, and join-the-dot conditions; (d) measured participants' performance on 12 DRM word lists—counterbalanced for BAS—rather than on a single list; and (e) measured both free-recall and recognition performance.

This experiment was conducted in a double-blind randomized controlled trial (Gilder & Heerey, 2018). Our hypotheses were as follows:

In the free-recall task, correct recall should be highest in the mindfulness condition and lowest in the mind-wandering condition. False recall should be highest in the mind-wandering condition and lowest in the mindfulness condition.[1] In the recognition task, (a) correct recognition should be highest in the mindfulness condition and lowest in the mind-wandering condition, (b) false recognition should be highest in

the mind-wandering condition and lowest in the mindfulness condition, and (c) filler recognition should be highest in the mind-wandering condition and lowest in the mindfulness condition.

On the recognition task, (a) "remember" responses for correct recognition should be highest in the mindfulness condition and lowest in the mind-wandering condition, (b) "remember" responses for false recognition should be highest in the no-manipulation condition and lowest in the mindfulness condition, and (c) "remember" responses for filler recognition should be highest in the no-manipulation condition and lowest in the mindfulness condition. Further, (a) "know" responses for correct recognition should be highest in the mindfulness condition and lowest in the mind-wandering condition, (b) "know" responses for false recognition should be highest in the mind-wandering condition and lowest in the mindfulness condition, and (c) "know" responses for filler recognition should be highest in the mind-wandering condition and lowest in the mindfulness condition.

Method

Participants

Our target sample size was informed by the effect size from Wilson et al. (2015). The effect size (Cohen's *d*) in their Study 1 was 0.5 (a medium-sized effect). The equivalent effect size (Cohen's *f*) in an analysis of variance (ANOVA) is 0.25. To power our experiment at 95% with the same effect size (0.25) as Wilson et al.'s, we required at least 250 participants. This differs from the power-analysis calculation in our preregistration, in which we incorrectly stated that the power analysis suggested a total sample size of 280. Hence, our final sample size provided more power than planned.

A total of 302 participants were recruited through Keele University's School of Psychology research-participation scheme, through social media, and through fliers posted on the Keele University campus. Participants either received course credit or were paid £7 for participating. Participants were at least 18 years of age ($M = 23.44$ years, $SD = 9.80$), and English was their first language.

Following our preregistered exclusion plan, we removed data from 15 participants: 6 were nonnative English speakers (violating one of our eligibility criteria); 2 overran the 1-hr time slot that the participants were booked for, so their participation had to be terminated early; 4 did not complete the experiment; 1 did not provide any free-recall data; and 2 completed free recall only after some of the lists, thus violating the stopping rule that required complete sets of data.

This left us with 287 participants. A sensitivity analysis (see the Supplemental Material available online) showed that our final sample size gave us 80% power to detect an effect size (f) as small as 0.185, which is 25% smaller than the effect size reported by Wilson et al. (2015).

Materials and procedure

The study received ethical approval from the Keele University Ethical Review Panel on May 10, 2017 (Document Number ERP1331). Participants were tested individually in a lab with an experimenter present to ensure full participation (e.g., that the participants were not using their mobile phones). The experimenter was blind to the experimental condition of each participant, as random assignment to condition was done using Qualtrics software (https://www.qualtrics.com/). After providing informed consent, participants completed the State Mindfulness Scale (SMS; Tanay & Bernstein, 2013), which consists of two subscales—a 15-item state mindfulness-of-mind scale (SMS mind) and a 6-item state mindfulness-of-body scale (SMS body). The 21 items were presented in random order. Participants then completed the relevant mindfulness, mind-wandering, or control-condition activity.

Wilson et al. (2015) used 15-min mindfulness and mind-wandering inductions recorded by Marilee Bresciani Ludvik at the Rushing to Yoga Foundation. When we requested these recordings from Wilson, he informed us that since the principal investigator, Edmund Fantino, had recently passed away, the precise recordings were not available (Wilson, personal communication, March 3, 2016). He provided similar recordings by the same person, so we used these instead, with participants listening to them via headphones. In the control condition, participants were asked to complete paper-based join-the-dot puzzles for 15 min (this task was identified by Friese, Messner, & Schaffner, 2012, as being "neither boring nor resource demanding"; p. 1019).

Participants then completed the SMS items in a different random order and the Retrospective Mind-Wandering Scale (the Thinking Content component of the Dundee Stress State Questionnaire; Matthews et al., 2013), which consists of an eight-item task-related-interference (TRI) scale and an eight-item task-unrelated-thought (TUT) scale.

Eighteen lists of 15 words were selected from Roediger et al. (2001). Each participant saw 12 of the 18 lists, and the lists were counterbalanced by dividing them into three sets. The lists were chosen and counterbalanced on the basis of the two factors that predict false recall—BAS and veridical recall—and also on the norms for

false recall and false recognition for each list. The 15 words per list were presented individually in the Qualtrics default black font, size 36, for 1.5 s in the middle of the screen. After each list was presented, participants were given 3 min to type as many words as they could remember from the list. Once this was repeated for all 12 lists, participants completed a remember/know/guess recognition test. The recognition test consisted of 72 items: 36 presented items (3 from each list), 12 critical lure items (1 from each list), and 24 filler items (3 list items and 1 lure item from the 6 nonpresented lists, which were counterbalanced across participants). For each item, participants had to identify whether it was old or new, and then for those items identified as old, they had to select from among "remember," "know," or "guess" responses. A "remember" response indicated that a participant was able to consciously recollect that the item appeared in the original list, whereas a "know" response indicated that the recognition was based on a feeling of familiarity for the item in that context. The definitions (adapted from Dewhurst & Anderson, 1999) were provided in the instructions and again every time participants had to make a selection.

Results

The analytical approach we used was to present standard null-hypothesis significance tests (NHSTs) together with Bayesian analysis (in the form of Bayes factors). The analysis consisted of a series of one-way between-participants ANOVAs. For all tests, the independent categorical variable was state of mind (three levels: mindfulness vs. mind wandering vs. join-the-dots).

For the NHSTs, omnibus ANOVAs were followed by Tukey's honestly significant difference (HSD) pairwise comparisons with the criterion for significance (α) set at .05. Although we present all of these tests for completeness, we interpret the HSD tests only when there was a significant omnibus ANOVA. For the Bayesian analyses, we used default Bayes factor tests for ANOVA designs (Rouder, Morey, Verhagen, Swagman, & Wagenmakers, 2017) using the *BayesFactor* package for R (Version 0.9.12-4.2; Morey & Rouder, 2018).

We conducted three model comparisons using Bayes factors. The first model comparison was a null model (i.e., all three levels of the design are equal) against a full model (i.e., all three levels are not equal), denoted by $BF_{null–full}$. This comparison allowed quantification of the degree of support for a model showing some effect versus a model showing no effect. Another model—the order-restricted model—was then constructed. In contrast to the full (unrestricted) model, in which all levels of the design are assumed to be different, order-restricted models test whether the data fit a predicted ordering of the factor-level effects (e.g., the mindfulness

score is greater than the join-the-dots score, which in turn is greater than the mind-wandering score). In a second model comparison, then, this order-restricted model was compared with the full (unrestricted) model; this comparison—denoted by $BF_{restricted–full}$—allowed quantification of the degree of support for a model showing a specifically ordered (and predicted) effect versus a model showing some (unrestricted) effect. Thus, in the presence of an effect, this model comparison allowed us to test whether the ordering of the factor levels matched our preregistered hypotheses. The third model comparison—$BF_{restricted–null}$—compared the order-restricted model with the null model. This model comparison allowed us to compare a null model with a model capturing our preregistered hypotheses. We followed the recommendations set out by Morey (2015) for testing the order-restricted models using the Bayes factor package.

Note that with a Bayes factor for model comparison between model *X* and model *Y*—denoted by $BF_{modelX–modelY}$—the evidence in favor of model *X* is given by $BF_{modelX–modelY}$ itself, whereas the evidence for model *Y* is given as the inverse of this (i.e., $BF_{modelX–modelY} = 1/BF_{modelX–modelY}$). The reader should take note of the ordering of the subscript of Bayes factor reporting to note which model the data are providing support for. We interpreted Bayes factors between 1 and 3 as anecdotal evidence, between 3 and 10 as moderate evidence, between 10 and 30 as strong evidence, between 30 and 100 as very strong evidence, and greater than 100 as extremely strong evidence (Lee & Wagenmakers, 2013).

Manipulation checks

Before turning to the main analysis, we wanted to ascertain that our manipulations of mindfulness and mind wandering worked by assessing their impact on SMS, TRI, and TUT scores. For the SMS scale and its components, we used difference scores as the dependent variable by (postmanipulation score minus premanipulation score). The descriptive statistics for the manipulation checks are in Table 1. The top left panel of Figure 1 shows standardized effect sizes of between-condition comparisons for all scales.

For SMS total score (i.e., SMS mind and SMS body scores combined), there was an effect of state of mind, $F(2, 281) = 27.11$, $p < .001$, $\eta^2 = .16$, 95% confidence interval (CI) = [.09, .24]; the Bayes factor supported the presence of an effect (i.e., the full model) compared with the null model, $BF_{full–null} = 5.01 \times 10^8$. Note that our planned post hoc Bayesian model comparisons could not be conducted on these data, because the data generally did not conform to our order-restricted predictions. Often there was no significant difference

Table 1. Mean Values for the Manipulation Checks in All Three Conditions

Measure	Condition		
	Mindfulness	Mind wandering	Join the dots
SMS total	15.8 (1.52)	12.8 (1.46)	0.11 (1.77)
SMS mind	11.1 (1.15)	12.0 (1.07)	1.6 (1.32)
SMS body	4.66 (0.50)	0.79 (0.60)	−1.49 (0.64)
Task-related interference	18.2 (0.76)	18.1 (0.57)	24.3 (0.64)
Task-unrelated thought	19.0 (0.82)	24.1 (0.68)	15.7 (0.74)

Note: Standard errors of the mean are given in parentheses. The total State Mindfulness Scale (SMS) consists of 21 items scored on a 5-point scale; scores were also calculated for the two subscales: state mindfulness of mind (15 items) and state mindfulness of body (6 items). Means for the SMS measures are difference scores (postmanipulation score minus premanipulation score). The task-related-interference scale consists of 8 items scored on a 5-point scale, and the task-unrelated-thought scale consists of 8 items scored on a 5-point scale.

between scores in the mindfulness and mind-wandering conditions, yet our predictions contained the constraint that they would differ. Consequently, no samples from the posterior distribution conformed with our predicted ordering of the factor levels, and thus a value of zero was entered in the denominator of the Bayes factor test, producing a division including zero (which cannot be calculated meaningfully without incurring infinities). However, as a reviewer observed, entering such a model that is clearly wrong into formal model comparisons is likely not very informative. Tukey's HSD tests showed no significant difference between scores in the mindfulness and mind-wandering conditions ($p = .394$), but scores in both the mindfulness ($p < .001$) and mind-wandering conditions ($p < .001$) were greater than those in the join-the-dots condition.

A similar pattern of results was found for the SMS mind scale, $F(2, 281) = 23.90$, $p < .001$, $\eta^2 = .15$, 95% CI = [.07, .22], $BF_{full-null} = 3.71 \times 10^7$, in which there was no significant difference between scores in the mindfulness and mind-wandering conditions ($p = .843$), but both scores in the mindfulness ($p < .001$) and mind-wandering conditions ($p < .001$) were greater than those in the join-the-dots condition. For the SMS body scale, there was again an effect of state of mind, $F(2, 281) = 27.81$, $p < .001$, $\eta^2 = .17$, 95% CI = [.09, .24], $BF_{full-null} = 8.72 \times 10^8$, but all comparisons were now significant: Scores in the mindfulness condition were greater than those in the mind-wandering condition ($p < .001$), scores in the mindfulness condition were greater than those in the join-the-dots condition ($p < .001$), and scores in the mind-wandering condition were also greater than those in the join-the-dots condition ($p = .016$).

For the TRI questionnaire, there was a significant effect of state of mind, $F(2, 281) = 27.54$, $p < .001$, $\eta^2 = .16$, 95% CI = [.09, .24], $BF_{full-null} = 7.07 \times 10^8$. There was no significant difference between scores in the

mindfulness and mind-wandering conditions ($p = .987$), but scores in the mindfulness condition were lower than those in the join-the-dots condition ($p < .001$), and mind-wandering was also lower than join-the-dots ($p < .001$). TUT scores also exhibited a significant effect of state of mind, $F(2, 281) = 35.18$, $p < .001$, $\eta^2 = .20$, 95% CI = [.12, .28], $BF_{full-null} = 2.99 \times 10^{11}$. Scores in the mindfulness condition were significantly lower than those in the mind-wandering condition ($p < .001$), scores in the mindfulness condition were greater than those in the join-the-dots condition ($p = .003$), and scores in the mind-wandering condition were also greater than those in the join-the-dots condition ($p < .001$).

Effects on memory

In this section, we present the results on the effects of state of mind on correct and false memory for both recognition and recall data. The descriptive statistics for all tests are shown in Table 2. For ease of exposition, we present the ANOVA results for all tests in Table 3 and the results of the Bayesian model comparisons in Table 4. Plots of standardized effect sizes for all between-condition comparisons can be seen in Figure 1.

For the recognition data, the NHST analysis showed no significant effects of state of mind on any of the measures of correct or false memories (lowest $p = .077$). For all of these tests, the Bayesian model comparison $BF_{null-full}$ favored the null model: For the measure of total correct recognition, the evidence in favor of the null model compared with the full model was only anecdotal; for all other measures, the $BF_{null-full}$ was either moderate ("remember," correct recognition; "know," false recognition) or strong (all others). We also found that the null model was preferred against our order-restricted models—that is, $BF_{null-restricted}$—in all
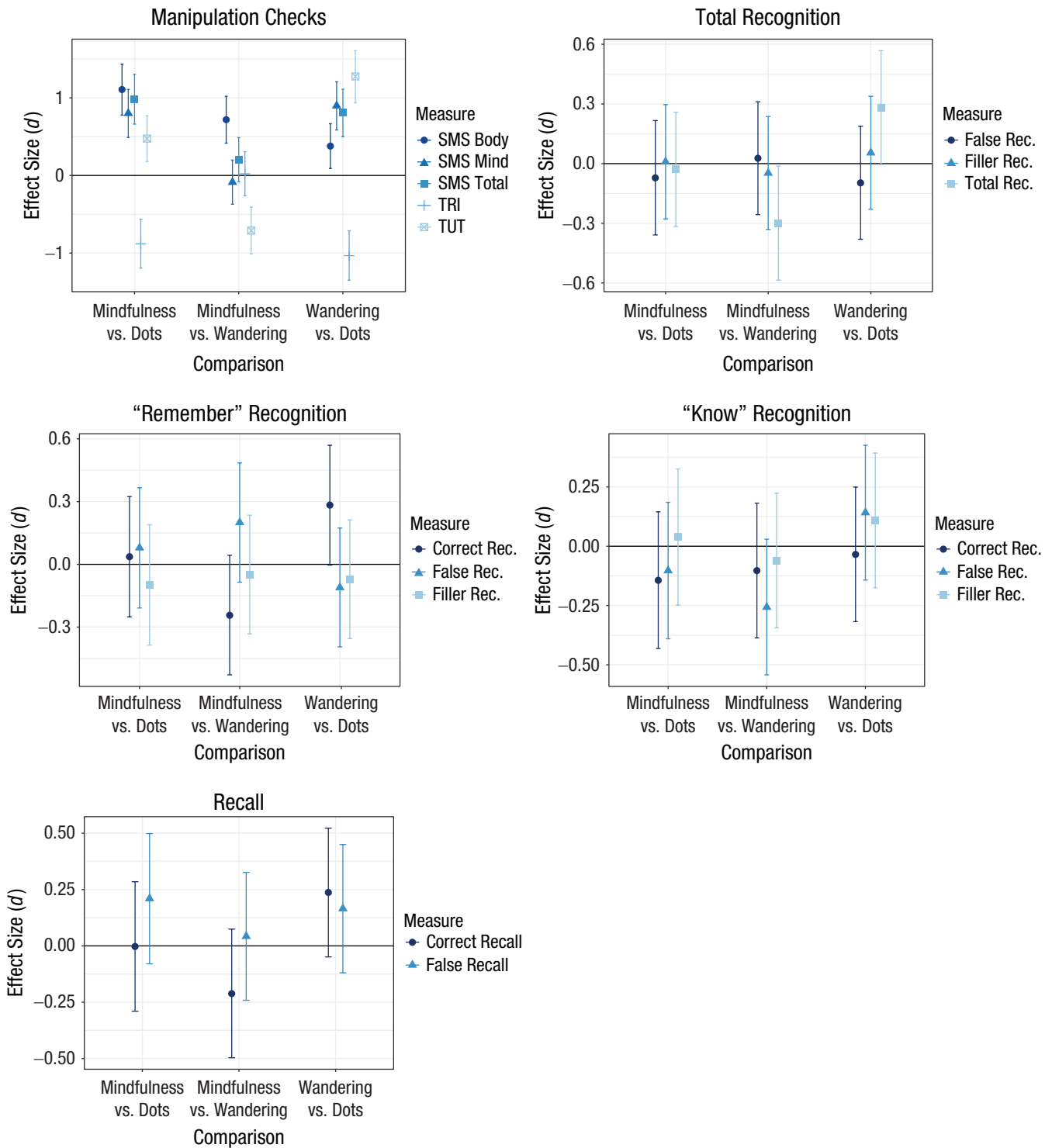
Fig. 1. Standardized effect-size estimates (Cohen's *d*s) for all between-conditions pairwise comparisons among the mindfulness, join-the-dots, and mind-wandering conditions. Results are shown separately for each manipulation check (top left) and for all measures of memory performance. Three measures were assessed for the State Mindfulness Scale (SMS): the state mindfulness-of-mind subscale, the state mindfulness-of-body subscale, and the total SMS. Error bars indicate 95% confidence intervals. TRI = task-related interference; TUT = task-unrelated thought; Rec. = recognition.

cases; for "remember" (filler) and "know" (false recognition) items, the evidence for this preference was moderate, but it was strong in all other cases. Thus, in

summary, we found no evidence for an effect of state of mind on any measures of correct- or false-recognition memory.

Table 2. Mean Proportions for All Dependent Variables in All Three Conditions

Dependent variable	Condition		
	Mindfulness	Mind wandering	Join the dots
Total correct recognition	.74 (.02)	.78 (.01)	.74 (.54)
Total false recognition	.73 (.02)	.73 (.02)	.75 (.02)
Total filler	.12 (.01)	.13 (.01)	.12 (.02)
"Remember" correct recognition	.50 (.02)	.54 (.02)	.49 (.02)
"Remember" false recognition	.40 (.03)	.35 (.03)	.38 (.03)
"Remember" filler	.34 (.11)	.40 (.11)	.54 (.26)
"Know" correct recognition	.17 (.01)	.18 (.01)	.18 (.01)
"Know" false recognition	.25 (.02)	.29 (.02)	.27 (.02)
"Know" filler	.05 (.01)	.05 (.01)	.04 (.01)
Correct recall	.51 (.01)	.53 (.01)	.51 (.01)
False recall	.37 (.02)	.37 (.02)	.33 (.02)

Note: Standard errors are given in parentheses. Correct recognition and correct recall refer to the correct recognition of list items. False recognition and false recall refer to the false recognition of critical lures.

For the free-recall data, a similar picture emerged. The NHST analysis showed no significant effect of state of mind for either correct or false recall. The Bayesian model comparison $BF_{null–full}$ provided moderate support for the null model in both cases. For the order-restricted tests, the correct-recall data were better predicted by

Table 3. Results of Null-Hypothesis Significance Tests of the Effect of State of Mind on Measures of Memory

Memory measure	Omnibus ANOVA	Pairwise comparison[a]		
		Mindfulness vs. mind wandering	Mindfulness vs. join the dots	Mind wandering vs. join the dots
Total correct recognition	$F(2, 281) = 2.58$, $p = .077$, $\eta^2 = .02$, 95% CI = [0, .06]	.101	.977	.156
Total false recognition	$F(2, 281) = 0.24$, $p = .789$, $\eta^2 < .01$, 95% CI = [< .01, .02]	.980	.883	.781
Total filler	$F(2, 281) = 0.08$, $p = .923$, $\eta^2 < .01$, 95% CI = [< .01, .001]	.949	.998	.926
"Remember" correct recognition	$F(2, 281) = 2.20$, $p = .113$, $\eta^2 = .02$, 95% CI = [0, .05]	.220	.965	.134
"Remember" false recognition	$F(2, 281) = 0.89$, $p = .412$, $\eta^2 < .01$, 95% CI = [< .01, .03]	.381	.847	.722
"Remember" filler	$F(2, 281) = 0.31$, $p = .733$, $\eta^2 < .01$, 95% CI = [< .01, .02]	.975	.726	.843
"Know" correct recognition	$F(2, 281) = 0.46$, $p = .630$, $\eta^2 < .01$, 95% CI = [< .001, .02]	.769	.623	.966
"Know" false recognition	$F(2, 281) = 1.49$, $p = .226$, $\eta^2 = .01$, 95% CI = [0, .04]	.206	.776	.560
"Know" filler recognition	$F(2, 281) = 0.24$, $p = .784$, $\eta^2 = .15$, 95% CI = [0, .02]	.916	.954	.767
Correct recall	$F(2, 281) = 1.60$, $p = .203$, $\eta^2 = .01$, 95% CI = [0, .04]	.269	.999	.276
False recall	$F(2, 281) = 1.13$, $p = .325$, $\eta^2 = .01$, 95% CI = [0, .04]	.954	.332	.487

Note: Values in brackets are 95% confidence intervals for the η^2 effect-size estimates. Post hoc comparisons are interpretable only in the event of a significant omnibus analysis-of-variance (ANOVA) result. Correct recognition refers to the correct recognition of list items. False recognition refers to the false recognition of critical lures. CI = confidence interval.
[a]Numbers in these columns are Tukey's honestly significant difference (HSD) p values.

Table 4. Bayes Factors for All Model Comparisons for the Different Memory Measures

	Model comparison		
Memory measure	Null vs. full	Full vs. restricted	Null vs. restricted
Total correct recognition	2.56	62.50	166.67
Total false recognition	21.28	1.47	31.25
Total filler	24.39	0.75	18.52
"Remember" correct recognition	3.62	31.25	111.11
"Remember" false recognition	11.76	3.75	43.48
"Remember" filler	20.00	0.49	9.71
"Know" correct recognition	17.54	3.07	52.63
"Know" false recognition	6.85	0.68	4.63
"Know" filler recognition	21.28	0.75	15.87
Correct recall	6.21	0.38	2.34
False recall	9.52	1.64	15.62

Note: In a model comparison of model *X* vs. model *Y*, a Bayes factor larger than 1 indicates support for model *X*, and a Bayes factor lower than 1 indicates support for model *Y*.

the null model but only at anecdotal levels, $BF_{null–restricted} = 2.34$. The null model was a much better predictor of the data than the restricted model, $BF_{null–restricted} = 15.62$, which is strong evidence in favor of the null model.[2]

Discussion

In summary, the state-of-mind inductions worked: The mindfulness induction induced mindfulness, and the mind-wandering induction induced mind wandering. However, there was no evidence of a difference in the levels of either correct or false memory for recall or recognition among the mindfulness, mind-wandering, or join-the-dots conditions. Thus, none of our hypotheses were supported; furthermore, neither were the previous findings by Wilson et al. (2015, Experiment 1) or Rosenstreich (2016, Experiment 2), who found that mindfulness increased false memories. Instead, our findings are consistent with those of Baranski and Was (2017), who also found no evidence for a difference in either true or false recall or recognition-memory performance between mindfulness and mind-wandering conditions in their first experiment or in false recall between the conditions in their second experiment, which included a join-the-dots condition. One explanation for the discrepant findings across the five experiments is that the brief inductions used (all 15 min in length, except for Rosenstreich's Experiment 2, which was 30 min) were not sufficient to consistently induce the relevant state of mind to last throughout the subsequent tasks. Possibly the inductions are not long

enough, or one-off brief mindfulness manipulations—unless used with experienced meditators—may just not increase mindfulness. Furthermore, we used a double-blind procedure in which the experimenters did not know which condition participants were assigned to. Gilder and Heerey (2018) demonstrated the impact of a non-double-blind procedure on performance. It is unclear whether Wilson et al. used such a procedure in their Experiment 1 (they did in Experiment 2), whether Baranski and Was (2017) or Rosenstreich (2016) did, and whether such variance in procedures had any impact.

One key methodological difference between our study and previous research was our use of manipulation checks to ensure that the manipulations induced the state of mind they were claimed to. Despite the effectiveness of the inductions, there are additional findings to consider. First, not only did the mindfulness manipulation induce mindfulness, but also so did the mind-wandering manipulation, with higher postinduction scores for the total SMS and the SMS mind subscale. The join-the-dots condition, by comparison, did not induce mindfulness on any of the scales. Second, the mindfulness induction also induced mind wandering, leading to higher postinduction scores for both TRI and TUT. This is perhaps not surprising given that this brief mindfulness induction was likely the first exposure to mindfulness for many participants, and mind wandering is more prevalent in novice meditators (Lutz, Slagter, Dunne, & Davidson, 2008). The join-the-dots condition also induced mind wandering, although only on the

TRI and not on the TUT scale. Third, there was no significant difference between the mindfulness and mind-wandering inductions on the total SMS, the SMS mind subscale, or the TRI scale. Previous research contrasting mindfulness and mind wandering (e.g., Mrazek, Smallwood, & Schooler, 2012) has focused on dispositional mindfulness, whereas brief mindfulness inductions induce state mindfulness. It is possible that mind wandering might not be the appropriate control condition for state-mindfulness studies given the likely use of novice meditators. It may instead be that join the dots, or a similar activity, might be able to differentiate more clearly between the components at play during state mindfulness induced through a brief induction, because join the dots increased mind wandering but not mindfulness.

There are two potential problems with measuring states of mind: First, it is possible that demand characteristics distort the measurements; second, it is possible that the manipulation may wear off by the time the questionnaires are completed and the DRM lists are presented. However, as outlined in the introduction, it is not very satisfactory to assume that brief manipulations are sufficient to induce mindfulness or mind wandering, and we would further posit that there is not yet sufficient evidence to indicate that brief mindfulness and mind-wandering instructions do activate different states of mind. Further research is needed to address the longevity and nature of the states of mind induced by brief manipulations.

To conclude, more research is needed into the best control condition to use for state-mindfulness research. Our results are consistent with those of Baranski and Was (2017), showing no evidence for a difference in false-memory susceptibility among mindfulness, mind-wandering, and join-the-dots conditions. This suggests that it is too soon to say that "mindfulness plays havoc with memory" (Knapton, 2015).

Transparency

ORCID iD

Susan M. Sherman https://orcid.org/0000-0001-6708-3398

Acknowledgments

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620929302

Notes

1. In hindsight, we should have hypothesized that false memory would be reduced in both the mindfulness and mind-wandering conditions relative to the join-the-dots condition, but for different underlying reasons. In the mindfulness condition, false memory would be reduced because better source monitoring should be possible, and in the mind-wandering condition, it would be reduced because if the list items are not well encoded, then the spreading activation needed for false memories should not take place.
2. In Section C in the Supplemental Material, we present exploratory analyses assessing the extent to which individuals' scores on the state-of-mind measures predicted their memory performance irrespective of condition.

References

Baranski, M., & Was, C. A. (2017). Mindfulness meditation may not increase false-memory and may instead protect from false-memory susceptibility. *Mindfulness*, *8*, 1569–1579.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*, 822–848.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Dewhurst, S. A., & Anderson, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition*, *27*, 665–673.

Friese, M., Messner, C., & Schaffner, Y. (2012). Mindfulness meditation counteracts self-control depletion. *Consciousness and Cognition*, *21*, 1016–1022.

Gilder, T. S., & Heerey, E. A. (2018). The role of experimenter belief in social priming. *Psychological Science*, *29*, 403–417.

Knapton, S. (2015, September 9). How mindfulness plays havoc with memory. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/science/science-news/11855103/How-mindfulness-plays-havoc-with-memory.html

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge, England: Cambridge University Press.

Lutz, A., Slagter, H. A., Dunne, J. D., & Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in Cognitive Sciences*, *12*, 163–169.

Matthews, G., Szalma, J., Panganiban, A. R., Neubauer, C., & Warm, J. S. (2013). Profiling task stress with the Dundee Stress State Questionnaire. In L. Cavalcanti & S. Azevedo (Eds.), *Psychology of stress: New research* (pp. 49–91). Hauppage, NY: Nova Science.

Morey, R. D. (2015, January 18). Multiple comparisons with BayesFactor, Part 2 - order restrictions [Blog post]. Retrieved from http://bayesfactor.blogspot.com/2015/01/multiple-comparisons-with-bayesfactor-2.html

Morey, R. D., & Rouder, J. N. (2018). Package 'BayesFactor.' Retrieved from https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf

Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, *24*, 776–781.

Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: Finding convergence through opposing constructs. *Emotion*, *12*, 442–448. doi:10.1037/a0026678

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*, 1–75.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind-wandering and memory in a lecture. *Applied Cognitive Psychology*, *26*, 234–242.

Robinson, K. J., & Roediger, H. L., III. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*, 231–237. doi:10.1111/j.1467-9280.1997.tb00417.x

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814. doi:10.1037/0278-7393.21.4.803

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*, 385–407.

Rosenstreich, E. (2016). Mindfulness and false-memories: The impact of mindfulness practice on the DRM paradigm. *The Journal of Psychology*, *150*, 58–71. doi:10.1080/00223980.2015.1004298

Rouder, J. N., Morey, R. D., Verhagen, A. J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*, 304–321.

Tanay, G., & Bernstein, A. (2013). State Mindfulness Scale (SMS): Development and initial validation. *Psychological Assessment*, *25*, 1286–1299.

Wilson, B. M., Mickes, L., Stolarz-Fantino, S., Evrard, M., & Fantino, E. (2015). Increased false-memory susceptibility after mindfulness meditation. *Psychological Science*, *26*, 1567–1573. doi:10.1177/0956797615593705