“False Memory” is a linguistic convenience

Daniel M. Bernstein[1], Alan Scoboria[2], Lecia Desjarlais[1], and Kendall Soucie[2]

[1]Kwantlen Polytechnic University; [2]University of Windsor

In Press, Psychology of Consciousness: Theory, Research and Practice (December 23, 2017). This is not the copy of final record.

Abstract

The term, *false memory*, describes outcomes to various procedures and techniques, such as coming to believe that suggested false events occurred, acceptance of post-event misinformation, and recognition of critical lures in the Deese-Roediger-McDermott (DRM) procedure. The literature to date indicates that these memory errors inter-correlate poorly, if at all. However, issues with sample size and measurement reliability may have affected prior estimates. This study examined links among these three techniques in a sizeable sample (*N* = 360), with attention to the reliability of procedures and measures. Weak but statistically reliable correlations emerged between false event suggestion and post-event misinformation, and between post-event misinformation and DRM intrusions; correlations between false event suggestion and DRM were weak and inconsistent. These effects are influenced by largely independent underlying mechanisms. The term, 'false memory' lacks precision, and requires qualification.

Introduction

Consider the following: Your relative confuses you with your sibling, and tells you erroneously that you got sick eating peach yogurt as a child. At first, you don't recall this experience, but over time, you come to believe that the event happened in your childhood. We will refer to this type of memory error as the *false event suggestion effect* (Loftus & Pickrell, 1995). Now consider this: You witness a robbery in which the robber wore a light blue hoodie. Later, someone erroneously tells you that the robber wore a dark green hoodie. Eventually, you come to believe that the robber wore a dark green hoodie. We refer to this type of memory error as the *post-event misinformation effect* (Loftus, Miller, & Burns, 1977), or more simply, ‘misinformation effect'. Finally, consider this: You study a list of semantically-related words (e.g., bed, tired, rest, dream, pillow) and later recall a critical lure word that was not on the list (e.g., sleep). We will refer to this type of memory error as the *Deese-Roediger McDermott associative memory illusion effect* or more simply, the *DRM effect* (Deese, 1959; Roediger & McDermott, 1995). What do these memory errors have in common? The short answer is, not much. Although researchers refer to all these errors as *false memory*, there is little evidence that these errors relate to each other. We add to that dearth of evidence here by showing that false memories created via false event suggestion, misinformation and the DRM either correlate weakly or not at all with each other.

Researchers have used various techniques to study and elicit erroneous memory reports (see Pohl, 2017; Newman & Lindsay, 2009; Roediger, 1996). Our focus here is on the three techniques just described: false event suggestion, misinformation and the DRM. For simplicity, we use the same terms to denote both the technique used to elicit erroneous memory reports and the effects observed. To denote effects, we simply add “effect” to the term. For example, we use the terms, “DRM” and “DRM effect” to denote the technique and effect, respectively. Further, each false memory technique that we consider involves an elicitation technique and associated measures of false memory. We use the term, "technique," to represent the general method used to elicit false memory. We use the term, "measure," to represent the actual tool used to assess false memory.

Brief primer on false event suggestion, misinformation and DRM

First, some historical context. For most of the 20th century, memory researchers tended to focus on errors of omission—what people forgot to report about their experiences (often meaningless word stimuli, called consonant-vowel-consonant trigrams or CVCs). With the emergence of controversy surrounding repressed versus recovered memories in the early 1990s, much of the prior focus on errors of omission shifted to include errors of commission—or what many researchers refer to as false memory (see Bartlett, 1932; Kihlstrohm, 1998; Lindsay & Read, 1994; Loftus, 1993). Although we applaud this shift in focus, our study is a cautionary tale that all false memories are not created equal.

To study the false event suggestion effect, researchers suggest to participants that a false autobiographical event occurred in the participant's past. There are two broad types of studies in this category, false memory implantation studies and false food-event suggestion studies, which differ by the intensity of the suggestion and the number of techniques employed as part of the suggestive technique. False memory implantation studies involve attempts to examine whether people develop complete robust, or *rich* false memories that closely resemble typical memories for genuine events (Loftus & Bernstein, 2005). For example, researchers might falsely tell a participant named Suzie that they spoke with her parent who said that Suzie had taken a ride in a hot-air balloon as a child (Wade, Garry, Read, & Lindsay, 2002). Although Suzie never took such a ride, she tries to recall details from this event. When she cannot recall the event, Suzie might be asked to imagine how the event could have happened. Imagining an event often leads to small but reliable increases in belief that the event occurred (Garry, Manning, Loftus, & Sherman, 1996; Johnson, Foley, Suengas, & Raye, 1988). When asked again one week later about the event and following repeated efforts to retrieve a memory, Suzie may come to believe that the event occurred in her past. Consequently, Suzie may furnish a vivid description of the details about this event that never happened, and even claim that she has recalled a vivid memory. In such studies, a substantial minority of individuals (20% - 30%) come to report vivid false memories, and a majority of participants exhibit some degree of belief that the event occurred (Scoboria et al., 2017).

The second type of false suggestion technique examines the effects of briefer, less comprehensive suggestions on the development of false autobiographical belief and false memory. For example, Jim completes several measures related to food preferences, belief in the occurrence of food-related events, and other individual difference variables. One week later, Jim reads a computer printout allegedly based on his initial responses that indicates the likelihood that certain food-related events occurred during his childhood. The report includes information that Jim became sick after eating egg salad. He then immediately completes post-suggestion measures regarding preference for and intention to eat numerous foods (including the target item), belief that the event occurred during childhood, and a measure in which he indicates whether he has a vivid memory, believes that the event occurred without having a memory, or neither. While many of these types of studies have used suggestions about experiences with foods, other suggestions have been used too (e.g., Berkowitz et al., 2008; Bernstein & Loftus, 2009). As with false memory implantation studies, false food-event suggestion studies result in more participants reporting beliefs for the suggested event than memories for the suggested event. A combined analysis of eight false food-event suggestion studies indicates that the development of false belief and not the development of false memory leads to suggestion-consistent changes in related behavioral outcomes such as food preferences and intention to eat the target food (Bernstein, Scoboria & Arnold, 2015).

To study the misinformation effect, researchers use subtle, moderate or strong suggestion. After witnessing an event, participants answer a series of questions about the event. Researchers might subtly suggest a false detail about the event by the wording of a question: e.g., “how fast was the white sports car going as it passed the barn on the country road?” (Loftus, 1975). In this example, there might have been no barn on the road, but the subtle mention of a barn leads witnesses to remember falsely that there was a barn. Instead, researchers might use moderate or strong suggestion and falsely tell a participant who witnessed a recorded crime that she initially claimed that the suspect wore a hat when, in fact, the witness initially said nothing about a hat. Or researchers might tell the witness that another witness noted the suspect's hat. In these latter two cases, the witness often (upwards of 70% of participants, depending on question wording and procedure) comes to remember falsely that the suspect wore a hat (see Loftus, 2017; Pickrell, McDonald, Bernstein, & Loftus, 2017).

To study the DRM effect, researchers use semantically-associated word lists. They might ask participants to study a list of words, such as the example provided in our first paragraph, that all relate to a non-presented (critical lure) word. Participants then try to recall and/or recognize the words on the list. Participants very often come to recall and recognize the critical lure word (upwards of 80%, depending on word list, Stadler, Roediger, & McDermott, 1999).

False event suggestion, misinformation, and DRM techniques all yield robust effects (see Gallo, 2010; Scoboria et al., 2017; Wade, Garry, & Pezdek, in press). Given the robustness of these effects and the ample literature on false memory, researchers have wondered how well false memory effects correlate with each other. Mazzoni (2002) proposed that memory distortions that arise automatically without social influence are distinct from suggestion-dependent memory errors that involve social interventions. Gallo (2010) reviewed early work in this area, and distinguished between two kinds of false autobiographical memories for entire events. These include naturally-occurring (e.g., memory for O.J. Simpson trial over an 18-month period, Platt, Lacey, Iobst, & Finkelman, 1998) and lab-induced false autobiographical memories (e.g., memory that one was lost in the mall as a child, Loftus & Pickrell, 1995). As of 2010, there was solid evidence linking naturally-occurring false memory and DRM false memory (Gallo, 2010). Conversely, as of 2010, there was little evidence linking lab-induced false memories to false event suggestion, misinformation, and DRM. We focus on lab-induced false memories here (see Table 1).

False event suggestion and DRM false memory correlations

We know of three published studies attempting to correlate false event suggestion and DRM false memories (see Table 1; for related work, see Patihis et al., in press, Experiment 1; Wilkinson & Hyman, 1998). These studies used the implantation technique (Loftus & Pickrell, 1995) that we described prior to create false beliefs and memories in participants. We briefly review these studies. Our interest here is in the correlation between the propensity to develop a false memory in the false event suggestion technique and the propensity to choose critical lure words in the DRM technique.

Qin, Ogle, and Goodman (2008) used the implantation technique to implant in participants false beliefs and memories for one of two early childhood events (having a birthday party at a popular restaurant; going to the hospital for an injury). Participants also completed the DRM. The authors found no correlation between false event suggestion and DRM recognition false memory. Otgaar and Candell (2011) re-tested children from prior false-memory studies that had used the implantation technique. Half the children had and half the children had not developed a false memory for an entire event (e.g., almost choking on a candy). Children completed the DRM. The authors found no correlation between false event suggestion and DRM false memory. Otgaar, Verschuere, Meijer, and van Oorsouw (2012) also used the implantation technique on children. Children reported their memory for an actual event from their first day at school, and then a false event (hot-air balloon ride). Children received two interviews, spaced one week apart. Children also completed the DRM. Children who developed what the authors termed, “full false memory” for the hot-air balloon ride falsely recalled more critical lures on the DRM (M = .35) than children who developed no false memory (M = .21). This is the only significant effect published thus far comparing false event suggestion to DRM false memory. As others have noted, caution is warranted when interpreting this result (Patihis et al., in press). The sample size was small (total N = 45, full false memory group N = 13), and the significant correlation only emerged after the authors removed “partial false memories” from the false memory group.

Post-Event Misinformation and DRM[1]

We know of five published studies attempting to correlate post-event misinformation and DRM false memory (see Table 1). We briefly review these studies, followed by a meta-analysis of them. Our interest here is in the correlation between the propensity to accept misinformation in the misinformation technique and the propensity to report critical lure words in the DRM technique.

In the first and most highly-powered study comparing misinformation to DRM false memory, Zhu, Chen, Loftus, Lin, and Dong (2013) had 432 participants watch two photographic slideshows. Participants received misinformation via written narrative 30 minutes later, and 10 minutes after this, they completed a recognition test of their memory for the original slideshows. Participants also completed the DRM. The authors found a small but statistically significant correlation (r = .12) between misinformation and DRM false memory. Ost, Blank, Davies, Jones, Lambert and Salmon (2013) had participants watch two minutes of Closed Circuit TV footage depicting an armed robbery. Participants received misinformation via written narrative or discussion with a confederate after completing the DRM. Participants then freely recalled the TV footage. The authors found no correlation between misinformation and DRM false memory. More recently, three additional papers have noted a similar null correlation between misinformation and DRM false memory (Calvillo & Parong, 2016; Monds, Patterson, & Kemp, 2016; Patihis et al., in press).

We conducted a meta-analysis of the five studies that estimated the correlation between acceptance of misinformation and DRM critical lure recognition. We converted Pearson correlations (see Table 2) to Fisher's Z values, and calculated standard errors on Z. We then calculated combined values, weighting by sample size. Finally, we converted back to a correlation the combined Z value and confidence interval. The combined correlation across the five studies (total N = 943) was r = .05 (95% CI -.08, .19). This correlation is clearly small and non-significant.

---

[1] We know of no study attempting to correlate lab-induced false event suggestion and misinformation false memory.

In retrospect, all these studies appear to have been statistically underpowered. Assuming that the actual population correlation is $r = .12$ (the estimate obtained by Zhu et al., the largest study conducted to date), a minimum N = 542 would be needed to detect this correlation as significant (assuming alpha of .05 and beta of .80); a more stringent test with alpha of .01 requires N = 806 (Faul, Erdfelder, Lang, & Buchner, 2007).

A note on psychometrics

There are two important features to consider about psychological measures, such as those used to measure the effects of false memory techniques. One feature is reliability, or the extent to which one's measures produce stable and consistent results. The second feature is validity, or the extent to which one's measures measure what one hopes they measure. Reliability of measures limits the magnitude of correlation between those measures that one can hope to observe (Cronbach & Meehl, 1955). Reliability (especially internal consistency, or the extent to which different items in a scale correlate with each other) typically rises as a function of the number of items included; reliability is further influenced by other factors such as test length, the degree to which the measure is unidimensional, and item difficulty.

Many of the published studies investigating correlations among the three false memory techniques and their associated measures that we consider have used versions of the measures with unreported or suspect reliability. This is especially the case with false event suggestion, which typically employs a single item for a single event. While methods exist for establishing the reliability of single-item measures, they have not yet been applied to this area (e.g., Robins, Hendin, & Trzesniewski, 2016; Wanous, Reichers, & Hudy, 1997). Even the DRM, which typically has good reliability, can be compromised if investigators use too few word lists or too few words per word list (see Gallo, 2010).

A note on theory

At a purely conceptual level, the three techniques we discuss have been linked to different theoretical memory constructs. The false event suggestion effect has been linked to a variety of familiarity, mental simulation, and episodic memory mechanisms, whereas false food-event suggestion draws more heavily on feelings of familiarity (Bernstein et al., 2005). The misinformation effect has been linked to episodic memory, and more specifically to interference and integrating/blending theories (Pickrell et al., 2017), while the DRM effect has been linked mainly to associative and semantic memory (Collins & Loftus, 1975; see Gallo, 2010). All three effects considered here have also been linked to source monitoring errors (Johnson, Hashtroudi, & Lindsay, 1993), in which participants fail to monitor the source of their “memories” and “beliefs” (e.g., events, event details, or words might feel familiar to participants because of, for example, suggestion or semantic relatedness to semantically-related words in a list).

Given the overarching goal of producing false memories using these measurement techniques, the resulting false memories would be expected to inter-correlate. However, the lack of observed correlations in studies to date may relate to the fact that the memory content and the psychometric properties across the three techniques differ in meaningful ways. If the measurement concerns can be addressed, and if the notion of a common false memory construct holds, then the three measurement techniques should produce false memory effects that inter-correlate. The correlations may also be weak because the theoretical mechanisms across techniques differ and there is, in fact, no unitary construct that unites or explains different false memory effects. However, if there were meaningful correlations among tasks, it would suggest that there were such a "common false memory construct" and, of course, the challenge would be to understand and elaborate the construct.

A note on individual differences

As in many prior studies, we attempted to correlate false memory with individual difference measures (Hyman & Billings, 1998; see Gallo, 2010). We focused on several measures that have been found to correlate with false memory in prior research. Of these, we highlight the concept of dissociation, which refers to failure to integrate thoughts, feelings, and experiences into one’s consciousness (e.g., looking in a mirror and not recognizing oneself). Dissociation is perhaps the most consistent individual difference measure to correlate with different types of false memory, and it is the one measure that correlated most strongly with our false-memory measures. We return to this in the Discussion.

The present study

Despite the expectation that false event suggestion, misinformation, and DRM false memory effects should inter-correlate, the evidence to date is wanting. The published work on this question has involved at most two out of the three effects in a single experiment, and many of the studies suffer from lack of power (small sample sizes), and poor or unknown reliability of the measures used. We attempt to remedy these shortcomings by comparing false event suggestion, misinformation, and DRM false memory in a single, moderately-powered study that uses measures with moderate to high reliability.

Method

The Research Ethics Board at Kwantlen Polytechnic University approved this study.

Participants

Undergraduate students at Kwantlen Polytechnic University in British Columbia, Canada participated in exchange for course credit or $15 cash. The study was conducted every semester (Fall, Spring, and Summer) from Fall 2011 to Summer 2012. A total of 523 participants completed at least one session. Of these, the 360 who completed all three sessions were included in the analyses (76% female, 22% male, 2% no response; mean age 20.91 years (SD = 4.25), range 17 to 59 years). Although 42% of the final sample spoke English as a first language, all participants indicated that they spoke English "very" (93%) or "moderately" (7%) well.

Measures

False food-event suggestion technique. We adopted the procedure from Bernstein et al. (2005), using the target food, spoiled peach yogurt, from Scoboria, Mazzoni and Jarry (2008). In Session 1, participants completed measures related to childhood experiences and food preferences (see Table 3). We created a personalized profile for each participant following this session. As in prior work, to increase the number of participants exposed to the false suggestion, we assigned most of the participants to the experimental group (suggestion, n = 358) and the balance to the control group (n = 97) in Session 2.

In Session 2, participants received feedback allegedly based on the measures from Session 1. We told participants that we had compared their responses to a database of thousands of other responses and that a personalized profile had been developed. Participants reviewed a sample profile to familiarize themselves with reading the output. Participants next received a sealed envelope with their name on it which contained their profile that appeared to have resulted from a scientific computer output. The chart contained a variety of filler items that were the same for all participants (e.g., "liked eating chocolate cake"). Participants were asked to study their chart and to complete the accompanying questions to help improve the quality of the database reports.

The suggestion group received feedback that they became sick after eating spoiled peach yogurt as a child. These participants then attempted to recall getting sick due to spoiled peach yogurt, and to write about the event. The control group received no feedback about spoiled peach yogurt, but otherwise followed the same procedure as the experimental group. Participants subsequently completed the post-measures.

The key measures for the false food-event suggestion technique were change in food preference ratings (peach and yogurt scores), change on the Autobiographical Belief and Memory Questionnaire, peach yogurt scale, change on the Food History Inventory sick on dairy item, and the post-manipulation breakfast behavior and Memory or Belief ratings.

False food-event suggestion measures

Our false-food measures fall under two distinct categories: those related to (1) the formation of false beliefs and memories; and (2) consequences of false beliefs and memories on attitudes and behaviours.

Food Preference. This 66-item measure assesses food preferences on a 7-point Likert-type scale (1 = strong dislike; 7 = strong preference; Bernstein, Laney, Morris, & Loftus, 2005). We used this measure to assess preferences for the target (peach yogurt) and non-target foods. We created two scales for the target food using factor scores from an exploratory factor analysis of six target items: peach (items: peaches, peach yogurt, peach nectar; α = .79; control group test-retest *r* = .82) and yogurt (peach, strawberry, vanilla, and blueberry yogurt; α = .87; control group test-retest *r* = .88). We also created a four-item face valid control scale, fish (items: salmon, sushi, tuna fish, clam chowder; α = .83; control group test-retest *r* = .95). This measure relates to consequences of false beliefs and memories on attitudes and behaviours.

Life Events Inventory. This 20-item measure assesses the likelihood that certain events occurred during childhood (Monaghan, Robinson, & Dodge, 1979). Participants indicated whether events happened to them before age 10 using 8-point Likert-type scales (1 = definitely did not happen; 8 = definitely happened). We included this measure to enhance our credibility that we had collected enough information to support the personalized profiles. This measure relates to the formation of false beliefs and memories.

Autobiographical Belief and Memory Questionnaire. This measure assesses general plausibility, personal plausibility, autobiographical belief, and memory for events. Participants rated five events using these four rating categories (20 ratings in total), including the target item, sick after eating spoiled yogurt, on an 8-point Likert-type scale (Scoboria, Mazzoni, Kirsch, and Relyea, 2004). We created a two-item scale for the target event by averaging the personal plausibility and autobiographical belief items (α = .71; control group test-retest *r* = .38). This measure relates to the formation of false beliefs and memories.

Food History Inventory. This 24-item version of the Life Events Inventory is specific to childhood events involving food. Participants indicated the likelihood that a variety of food-related events occurred before the age of 10. The single target item relevant to the suggested event was "Got sick after eating spoiled dairy" (Bernstein, Laney, Morris, & Loftus, 2005). Reliability could not be estimated for this single-item measure. This measure relates to the formation of false beliefs and memories.

Breakfast Behaviour Questionnaire. This 16-item measure assesses intention to eat various foods (Bernstein et al., 2005). Participants rated how likely they were to pick each food as they strolled a breakfast buffet, using an 8-point Likert-type scale (1 = definitely no; 8 = definitely yes). The single target item was "peach yogurt." Reliability could not be estimated for this single-item measure. This measure relates to consequences of false beliefs and memories on attitudes and behaviours.

Memory or Belief Questionnaire. This measure asks individuals to rate if they have a memory for an event, if they do not have a memory but believe that the event happened, or if they are positive that the event did not happen (Bernstein, Laney, Morris, & Loftus, 2005). We included the critical item, "got sick eating peach yogurt" and two distractor items, "spilled a bowl of punch at a wedding," and "were overweight for a significant portion of your childhood." This measure relates to the formation of false beliefs and memories.

Misinformation. We used the events developed by Okado and Stark (2005) and later adapted by Zhu et al. (2013). Participants watched two separate events presented as picture slideshows. One event involved a wallet being stolen from a woman. The second event involved a man breaking into a car. The order of the events was randomized. Each consisted of 50 picture slides that presented the event in sequence. Images were shown for 3500 ms, separated by blank white distractor slides displayed for 500 ms. Participants were instructed to watch the events carefully.

Participants then completed other measures for 30 minutes (see Table 3). Participants next viewed a slideshow that presented 50 sentences describing each event. For each event, the 50 sentences included 12 inaccurate descriptions (misinformation) and 38 accurate descriptions (consistent with the original slides). We told participants that they were reading an eyewitness account of the events and to pay careful attention. Participants were randomly allocated to one of two counterbalanced versions of the misinformation. For example, in the car break-in event, participants either read that the man used a hanger or they read that the man used a credit card to break into the car.

Participants completed other measures for 10 minutes, and then answered 36 forced-choice questions (18 per event; three response options per question) about the original slides. For the 24 critical questions, one response was correct, one was misinformation, and one was new (foil). Misinformed responses were foils for the counterbalanced group. The remaining 12 questions included the correct answer and two foils. Presentation of items was randomized.

The dependent variables for the misinformation procedure were number of correct, misinformed, and foil responses. Prior work has reported adequate reliability for this measure (Zhu et al., 2013). In our study, internal consistency for correct responses and total errors (the inverse of correct responses) was $\alpha = .60$, (36 items); for correct responses to misleading questions was $\alpha = .66$; and for errors to misleading questions $\alpha = .64$ (24 items). Reliability for foil responses was low ($\alpha = .28$), and therefore not examined further.

DRM. The DRM procedure facilitates recall and recognition of non-presented critical lures from word lists. Participants learned that they would study lists of words and that their memory would be tested. Participants heard a series of audio recordings of 30 lists consisting of 15 words per list read by a female researcher. We chose 30/36 lists from Stadler, Roediger, and McDermott (1999), retaining the lists whose critical lures elicited the 15 highest and 15 lowest levels of false recall (note that we omitted the “sleep” list, because it is the list most often used to describe the DRM technique to Psychology students). Many of the word lists have an average recall rate between 31% and 49%. We, therefore, dropped these additional lists with middling recall rates to ensure representation across the high and low ends of the validated word lists. After listening to each list, participants had 90 seconds to write as many words as they could recall. The recognition test immediately followed. Participants read a list of 90 words and circled as many as they recognized as presented in the word lists. The list included 30 critical lures, 30 words from the first position on each list, and 30 foils. This phase took approximately 1 hour.

The key dependent variables resulting from the DRM procedure were correct responses, critical lures, and intrusions (non-critical, non-presented words) during recall; and selection of correct responses, critical lures, and intrusions during recognition. Prior work has reported adequate reliability for this measure (see Gallo, 2010). In our study, internal consistency was as follows: recall correct α = .94; recall critical lures α = .76; recognition correct α = .83; recognition critical lures α = .85; recognition intrusions α = .73. Note that we could not calculate reliability for recall intrusions, because the number of these varies per person, and any two people may recall different words.

Individual difference measures

Vividness of Mental Imagery Questionnaire. This 16-item measure assesses the ability to visualize or form mental pictures. Prior work has reported adequate reliability for this measure (Marks, 1973). Four visual scenes appear, each with four prompts to imagine details. Participants rated the vividness of each detail on a 5-point Likert-type scale (1 = perfectly clear and vivid as normal vision; 5 = no image at all). We averaged items for a total score. Internal consistency for this measure was α = .92.

Dissociative Experiences Scale. This 28-item measure assesses trait dissociation—failure to integrate thoughts, feelings, and experiences into the stream of consciousness—on a 10-point Likert-type scale. Prior work has reported adequate discriminant validity for this measure (Bernstein & Putnam, 1986; Carlson & Putnam, 1993). We averaged items for a total score, and created three subscales. Internal consistency was as follows: overall measure α = .94; Amnesia (6 items) α = .82; Depersonalization (6 items) α = .82; Absorption (6 items) α = .85.

Persuadability. This 7-item measure assesses the tendency toward being persuaded by others. Participants rated items on a 7-point Likert-type scale (1 = totally disagree; 7 = totally agree. Prior work has reported adequate but rather low reliability for this measure (Kaptein, Lacroix, & Saini, 2010). We averaged items for a total score. Internal consistency for this measure was α = .68.

Gudjonsson Compliance Scale. This 20-item measure assesses the propensity to accept uncritically requests made by others. Prior work has reported adequate reliability for this measure (Gudjonsson, 1989). Participants rated items true or false. We summed true statements. Internal consistency for this measure was α = .76.

Creative Experiences Questionnaire. This 25-item measure assesses trait fantasy proneness. Prior work has reported adequate reliability and validity for this measure (Merckelback, Horselenberg, & Muris, 2001). Participants rated statements true or false. We summed true statements. Internal consistency for this measure was α = .73.

Procedure

Participants completed the three false memory techniques and their various measures (false food-event suggestion, misinformation, DRM), in addition to individual difference measures across three sessions. Table 3 outlines the study's overall procedure. Participants completed several additional questionnaires as filler tasks.

Session 1. The first session appeared as a series of questionnaires in a booklet labelled "Food and Beverage History Study." The measures included the pre-manipulation scales for the false food-event suggestion technique, and many of the individual difference measures. We instructed participants to complete the questionnaires in order. We collected demographic information at the end of the session. Session 1 took approximately 40 minutes to complete.

Session 2. The second session consisted of the first component of the misinformation technique (presenting events), the second phase (manipulation) of the false food-event suggestion technique, and the subsequent phases of the misinformation technique. As a cover story, we told participants that the study "examines how the brain stores memories for childhood events." Session 2 took approximately 60 minutes to complete.

Session 3. The third session contained the DRM technique, followed by debriefing. We told participants that this part of the study was about memory for word lists. Session 3 took approximately 90 minutes to complete.

Participants completed the study either individually (often the case in Session 1), or in small groups of up to six (in Sessions 2 and 3). Seven different research assistants worked on Sessions 1 and 3, though only one researcher tested any single group of participants. To maintain consistency, a single female research assistant administered Session 2. Participants completed the tasks in a fixed order within a small classroom.

Sessions followed one another weekly, resulting in a minimum of 3 days (Thursday to Monday), to a maximum of 10 days (Monday to Thursday) between sessions. The procedure was not conducted on Fridays. We omitted participants from analyses who did not complete the study in this time.

Results

Links among false food-event suggestion, misinformation, and DRM

The sample that we used to examine correlations with the false food-event suggestion variables included individuals whom we assigned to the suggestion group, and excluded individuals who rated either the 'sick on dairy' or 'belief sick peach yogurt' at the scale ceiling prior to the manipulation. Individuals who already had a strong belief for the target food event were not of interest in this study. See Appendix 1 and 2 for internal consistency and test-retest reliability of our false memory measures.

Correlations between false food-event suggestion and misinformation effects are in Table 4. Behavioral intention to eat peach yogurt and change in preference for peaches (decreased preference) correlated negatively with misinformation correct responses and correlated positively with misinformation mislead responses. We also conducted hierarchical regressions to control for pre-manipulation scores when estimating the correlations with change scores. The correlation between change in preference for peaches and misinformation correct responses decreased numerically from -.17

to -.15, but remained statistically below zero. The correlation between change in preference for peaches and misinformation mislead responses decreased from .15 to .14 but remained statistically above zero.

Correlations between false food-event suggestion and DRM effects are in Table 5. False food-event suggestion did not correlate with DRM recall, excepting a weak positive correlation between change in preference for yogurt and DRM correct recall. We conducted a hierarchical regression to control for pre-manipulation scores when estimating the correlation with change scores. The correlation decreased numerically from .15 to .13 and remained statistically above zero.

Two variables from the false food-event suggestion procedure correlated with DRM recognition. First, behavioral intention to eat peach yogurt correlated weakly and positively with DRM recognition intrusions. Second, and most notable here, dichotomous reports of belief in the suggested events on the Memory or Belief questionnaire following the manipulation correlated positively with DRM recognition critical lures.

Correlations between misinformation and DRM effects are in Table 6. We calculated Pearson correlations between the Misinformation and DRM techniques, including all participants who completed both techniques (*N* = 355). Based on 95% confidence intervals estimated on correlations, we found several weak but statistically significant correlations.

Misinformation correct responses correlated positively with DRM correct recall and correlated negatively with DRM recall intrusions, recognition critical lures and recognition intrusions. Misinformation mislead responses correlated negatively with DRM correct recall, and correlated positively with DRM recall intrusions, recognition critical lures, and recognition intrusions. Most notable here is the positive correlation between misinformation mislead responses and DRM recognition of critical lures.

Links between false memory and individual difference measures

Table 7 contains correlations among our three false memory techniques and individual difference measures. The most consistent results involved dissociation as measured by the Dissociative

Experiences Scale-II, which correlated weakly and positively with misinformation mislead responses, DRM recall and recognition intrusions, DRM recognition critical lures, and increase in belief for the suggested false food event.

Discussion

We conducted a three-part study over a three-week period to investigate links between different false memory measures. We chose three standard false memory measures: false event suggestion, misinformation, and the DRM. We also administered a battery of individual difference measures of dissociation, fantasy proneness, mental imagery, persuadability and compliance. Despite having the power to observe fairly small effects (N = 360; given alpha of .05 and beta of .80, this is sufficient to detect correlations of approximately .15 or higher), we found few and mostly weak correlations among our measures. These results add to the largely nonexistent to weak correlations among false memory measures in the literature, and suggest that the term, "false memory" is more of a linguistic convenience than a useful theoretical construct. We, therefore, conclude that false memory is a multifaceted construct with questionable utility (Newman & Lindsay, 2009; see also Hintzman, 2011).

Despite this sobering conclusion, the literature is rife with references to "false memory" as a unitary construct: "children and adolescents outperform adults on tests of cognitive biases, showing less of a susceptibility to…false memories" (e.g., Weldon, Corbin, Garavito, & Reyna, 2016, p. 71). The quoted statement is partially true in that younger children are less susceptible to DRM critical lures than are adults. However, other types of false memories show the reverse developmental pattern: younger children tend to be more susceptible to misinformation and suggestible memories than are adults (Brainerd & Reyna, 2005).

We are hardly the first to note the multifaceted nature of false memory (Brainerd & Reyna, 2005; Gallo, 2010; Ost et al., 2013; Roediger, 1996; Schacter, 2001; Wilkinson & Hyman, 1998). Nor are we the first to test for correlations among false event suggestion, misinformation, and the DRM (see Table 1). We are the first, however, to test all three techniques in a single, adequately-powered study with close attention to the reliability of measures. Several past studies have used insufficient sample sizes and/or measures with questionable reliability, thereby limiting the chances of detecting reliable correlations between measures (see Table 1; see Ost et al. for discussion of reliability of false memory measures). We chose our measures with reliability in mind, often necessitating a full hour of testing just to ensure the success of our measures (e.g., we administered 30 15-words DRM lists, elicited free recall after each list, and then administered a final recognition memory test). Even with these safeguards in place, we replicated the literature's weak correlations among these measures.

Given the lack of correlations, we must ask what processes underlie false memories arising from false event suggestion, misinformation, and the DRM. As we noted prior, these three techniques and the effects they produce relate to different theoretical memory constructs. Mazzoni (2002) proposed a model suggesting two different categories of memory distortion: (1) suggestion-dependent memory errors, and (2) errors as a product of memory reconstruction. Our study is partially compatible with this model. According to this model, we would expect false-event suggestion and the misinformation effect to correlate with each other and belong to category (1), and the DRM to belong to category (2). The low correlations between false event suggestion (suggestion-dependent memory error), misinformation (suggestion-dependent memory error), and the DRM (reconstruction memory error) demonstrate little overlap among these three false memory effects. The low correlation between misinformation and the DRM is compatible with Gallo's (2010) activation/monitoring framework: Misinformation activates bottom-up processes (feature overlap, feelings of familiarity etc.), while the DRM technique activates top-down processes (associations, schemas, etc.). A weak correlation between misinformation and DRM may indicate that individuals show impaired bottom-up and top-down source-monitoring activation overall.

In addition to our primary aim to explore correlations among three common false-memory measures, we explored correlations among false memory and individual differences. The latter included vividness of mental imagery, dissociation, persuadability, compliance, and creativity. The most consistent pattern of correlations emerged between dissociation and all three of our false-memory measures: Dissociation correlated positively with belief in our suggested food event (got sick eating peach yogurt as a child), misinformation mislead responses, and DRM recall and recognition intrusions, and DRM recognition critical lures. As others have noted, the likely mechanism explaining links between dissociation and susceptibility to forming false memories is reality monitoring failure (Gallo, 2010; Hyman & Billings, 1998; Johnson, 2006). On this account, people who fail to monitor their thoughts and feelings are more suggestible and, therefore, prone to form false memories (see Lynn et al., 2012; although see Dalenberg et al., 2012).

Limitations

We focused on three types of memory error: false event suggestion, misinformation, and the DRM. Therefore, we cannot say how any of these memory errors correlate with other memory errors, such as imagination inflation (Garry et al., 1996), false confabulation (Zaragoza et al., 2001), the revelation effect (Watkins & Peynirciouglu, 1990), or even naturally-occurring memory errors (Gallo, 2010).

As a reviewer of this study articulated, false memory is best conceptualized as a “useful category of memory effects - a set of phenomena in which subjects remember events that didn't actually happen to them. The occurrence of false memory is something to be explained, by invoking theoretical constructs...there's no reason that…underlying processes have anything to do with each other; or that false memories induced by their respective techniques will be correlated with each other.” We agree entirely with these statements, and urge memory researchers to heed these statements too.

An important limitation of this work pertains to the reliability of measures. Despite our attempts to ensure adequate reliability of all our measures, one of our measures by its nature has limited reliability. The false event memory measure and technique involve a single suggested event. We were able to form reliable measures for food preferences given the length of the measure. We were also able to construct an adequate two-item measure to assess the reliability of appraisals of the occurrence of the suggested event by combining the belief and personal plausibility items. Measures with fewer items tend to have poorer reliability, especially internal consistency. As we noted elsewhere, poor internal consistency limits correlations. Future work with this methodology might use validated multiple-item scales to assess belief in the occurrence of suggested events (Scoboria et al., 2014). The reliability of the single-item behavioral intention measure remains unknown. While it is possible to validate single-item measures, such efforts have not been undertaken in this area of research (Wanous, Reichers, & Hudy, 1997; Zimmerman et al., 2006). Future work on false memory should pay close attention to measurement issues, especially those related to reliability and validity. We note that the potential impact of scale reliability on estimated correlations can be estimated using the Spearman (1904) approach when planning future studies. Ost et al (2013) concluded that reliability of measurement was not a plausible explanation for the low magnitude of correlations between measures in their study. Given that the reliability coefficients in their paper are comparable or lower to those reported here, reliability of measurement is unlikely to have notably suppressed the relationships estimated here.

One limitation to the false event suggestion procedure is demand characteristics. The idea here is that participants intuit that we are attempting to create false memories. Once they intuit this, participants act in accordance with what they think that they should do to appear like "good" participants (Orne, 1962). As in prior studies that used this procedure, we attempted to disguise the true nature of the study (Bernstein et al., 2005; see Laney, Kaasa, Morris, Berkowitz, Bernstein, & Loftus, 2007). Although we do not have direct evidence arguing against demand characteristics, we do have indirect evidence. Our measure of compliance correlated modestly or not at all with our various false memory measures (see Table 7). Despite this indirect evidence, researchers should be mindful of demand characteristics in studies on false memory.

An overall limitation of this type of work is that it is laborious and time consuming, potentially deterring researchers. For example, it took over two years to complete data collection on this project, owing mostly to high attrition rates. The procedure required three visits, some lasting up to 90 minutes, spaced one week apart, and our participants were mostly university students who received course credit for their involvement. Unless greater incentives and safeguards are in place, researchers risk losing many participants through attrition. We urge researchers to look beyond the costs and to the benefits of conducting this kind of work in the future.

Conclusion

Three of the most commonly used measures of false memory—false event suggestion, misinformation, and the DRM—fail to correlate or correlate weakly with each other. Thus, performance on one type of false memory measure tells us little about performance on another false memory measure. The null to weak correlations that we observed, coupled with our large sample size and the adequate reliability of our measures, help explain the inconsistent findings in the literature. The best-used measures in the false memory literature do not correlate consistently. We, therefore, conclude that the term, "false memory" lacks precision. Researchers should use this term carefully, and should consider qualifying the term by the method used to elicit false memories.

Acknowledgments

We thank David Gallo and Roddy Roediger for helpful comments and suggestions regarding experimental materials and procedures, and we thank Eric Mah and two anonymous reviewers for comments on a previous draft of the manuscript. This work was supported by a Canada Research Chairs Award (950-228407) and Kwantlen Polytechnic University Chancellor's Chair Award to D. Bernstein, and funding from the Social Science and Humanities Research Council of Canada Insight Grant (435-2015-0040) to A. Scoboria.

References

Bartlett, F.C. (1932) *Remembering: A study in experimental and social psychology*, Cambridge: Cambridge University Press.

Berkowitz, S. R., Laney, C., Morris, E. K., Garry, M., & Loftus, E. F. (2008). Pluto behaving badly: False beliefs and their consequences. *The American Journal of Psychology, 121*(4), 643-660. doi:10.2307/20445490.

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False beliefs about fattening foods can have healthy consequences. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(39), 13724-13731. DOI:10.1073/pnas.0504869102

Bernstein, D. M., & Loftus, E. F. (2009). How to tell if a particular memory is true or false. *Perspectives on Psychological Science*, *4*(4), 370-374. DOI: 10.1111/j.1745-6924.2009.01140.x

Bernstein, D. M., Scoboria, A., & Arnold, R. (2015). The consequences of suggesting false childhood food events. *Acta Psychologica, 156,* 1-7. DOI: 10.1016/j.actpsy.2015.01.001

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, *174 (12)*, 727-735. DOI: 10.1097/00005053-198612000-00004

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. Oxford: Oxford University Press.

Calvillo, D. P., & Parong, J. A. (2016). The misinformation effect is unrelated to the DRM effect with and without a DRM warning. *Memory*, *24*(3), 324-333. DOI: 10.1080/09658211.2015.1005633

Carlson, E. B., & Putnam, F. W. (1993). An update on the dissociative experiences scale. *Dissociation, 6*(1), 16-27. DOI: 10.1023/A:1026275916184

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428. DOI: 10.1037/0033-295X.82.6.407

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. DOI: 10.1037/h0040957

Dalenberg, C. J., Brand, B. L., Gleaves, D. H., Dorahy, M. J., Loewenstein, R. J., Cardeña, E., . . . Spiegel, D. (2012). Evaluation of the evidence for the trauma and fantasy models of dissociation. *Psychological Bulletin*, *138*, 550–588. doi:10.1037/a0027447

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-22. DOI: 10.1037/h0046671

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. DOI: 10.3758/BF03193146

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, *38*(7), 833-848. DOI: 10.3758/MC.38.7.833

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, *3*(2), 208-214. DOI: 10.3758/BF03212420

Gudjonsson, G. H. (1989). Compliance in an interrogative situation: A new scale. *Personality and Individual Differences*, *10*(5), 535-540. DOI: 10.1016/0191-8869(89)90035-4

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth." *Perspectives on Psychological Science*, *6*(3), 253-271. DOI: 10.1177/1745691611406924

Hyman, I.E., Jr., & Billings, F.J. (1998). Individual differences and the creation of false childhood memories. *Memory, 6,* 1-20. DOI: 10.1080/741941598

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*(4), 371-376. DOI: 10.1037/0096-3445.117.4.371

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3-28. DOI: 10.1037/0033-2909.114.1.3

Kaptein M., Lacroix J., Saini P. (2010). Individual differences in persuadability in the health promotion domain. In: T. Ploug, P. Hasle, & H. Oinas-Kukkonen (Eds.) *Persuasive technology. 6137,* (94-105). Berlin: Springer. DOI: 10.1007/978-3-642-13226-1_11

Kihlstrom, J. F. (1998). Exhumed memory. In S. J. Lynn & K.M. McConkey (Eds.), *Truth in memory* (pages 3-31). New York: Guilford Press.

Laney, C., Kaasa, S.O., Berkowitz, S.R., Morris, E.K., Bernstein, D.M., & Loftus, E.F. (2008). The red herring technique: A methodological response to the problem of demand characteristics in false memory research. *Psychological Research*, 72, 362-375. DOI: 10.1007/s00426-007-0122-6

Lindsay, D.S., & Read, J.D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology*, 8, 281–338. DOI: 10.1002/acp.1780

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*(4), 560-572. DOI: 10.1016/0010-0285(75)90023-7.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist, 48*(5), 518-537. doi.org/10.1037/0003-066X.48.5.518.

Loftus, E. F. (2017) Eavesdropping on memory. *Annual Review of Psychology, 68*, 1-18. DOI: 10.1146/annurev-psych-010416-044138

Loftus, E. F., & Bernstein, D. M. (2005). Rich false memories: The royal road to success. In A. Healy (Ed.). *Experimental cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer* (pp. 101-113). Washington, DC, US: American Psychological Association.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(1), 19-31. DOI: 10.1037//0278-7393.4.1.19

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*(12), 720-725. DOI: 10.3928/0048-5713-19951201-07

Lynn, S. J., Lilienfeld, S. O., Merckelbach, H., Giesbrecht, T., McNally, R., Loftus, E.,... Malaktaris, A. (2014). The trauma model of dissociation: Inconvenient truths and stubborn fictions. Comment on Dalenberg et al. (2012). *Psychological Bulletin*, *140*, 896–910. doi:10.1037/ a0035570

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17-24. DOI: 10.1111/j.2044-8295.1973.tb01322.x

Mazzoni, G. (2002). Naturally occurring and suggestion-dependent memory distortions: The convergence of disparate research traditions. *European Psychologist*, *7*(1), 17-30. DOI: 10.1027//1016-9040.7.1.17

Merckelbach, H., Horselenberg, R., & Muris, P. (2001). The Creative Experiences Questionnaire (CEQ): A brief self-report measure of fantasy proneness. *Personality and Individual Differences*, *31*(6), 987-995. DOI: 10.1016/S0191-8869(00)00201-4

Monaghan, J. H., Robinson, J. O., & Dodge, J. A. (1979). The children's life events inventory. *Journal of Psychosomatic Research*, *23*(1), 63-68. DOI: 10.1016/0022-3999(79)90073-4

Monds, L. A., Paterson, H. M., & Kemp, R. I. (2016). Do emotional stimuli enhance or impede recall relative to neutral stimuli? An investigation of two “false memory” tasks. *Memory*, 1-8. DOI: 10.1080/09658211.2016.1237653

Newman, E. J., & Lindsay, D. S. (2009). False memories: What the hell are they for? *Applied Cognitive Psychology*, *23*(8), 1105-1121. DOI: 10.1002/acp.1613

Okado, Y., & Stark, C. E. (2005). Neural activity during encoding predicts false memories created by misinformation. *Learning & Memory*, *12*(1), 3-11. DOI: 10.1101/lm.87605

Orne, M.T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776-783. doi.org/10.1037/h0043424

Ost, J., Blank, H., Davies, J., Jones, G., Lambert, K., & Salmon, K. (2013). False memory ≠ false memory: DRM errors are unrelated to the misinformation effect. *PloS ONE*, *8*(4), e57939. DOI: 10.1371/journal.pone.0057939

Otgaar, H., & Candel, I. (2011). Children's false memories: Different false memory paradigms reveal different results. *Psychology, Crime & Law*, *17*(6), 513-528. DOI: 10.1080/10683160903373392

Otgaar, H., Verschuere, B., Meijer, E. H., & van Oorsouw, K. (2012). The origin of children's implanted false memories: Memory traces or compliance? *Acta Psychologica*, *139*, 397-403. DOI: 10.1016/j.actpsy.2012.01.002

Patihis, L., Frenda S. J., & Loftus, E. F. (in press). False memory tasks do not reliably predict other false memories. *Journal of Consciousness: Theory, Research, and Practice*.

Pickrell, J.E., McDonald, D.L.L., Bernstein, D.M., & Loftus, E.F. (2017). The misinformation effect. In R.F. Pohl (Ed.). *Cognitive illusions: Intriguing phenomena in thinking, judgment, and memory (2nd ed.).* Hove, UK: Psychology Press (pp.406-423).

Platt, R. D., Lacey, S. C., Iobst, A. D., & Finkelman, D. (1998). Absorption, dissociation, and fantasy-proneness as predictors of memory distortion in autobiographical and laboratory-generated memories. *Applied Cognitive Psychology*, *12*(7), 77-89. DOI: 10.1002/(SICI)1099-0720(199812)12:7<S77::AID-ACP601>3.0.CO;2-C

Pohl, R. F. (2017). *Cognitive illusions: Intriguing phenomena in judgement, thinking and memory*. New York: Routledge.

Qin, J., Ogle, C. M., & Goodman, G. S. (2008). Adults' memories of childhood: True and false reports. *Journal of Experimental Psychology: Applied*, *14*(4), 373-391. DOI: 10.1037/a0014309

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *27*(2), 151-161. DOI: 10.1177/0146167201272002

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35*(2), 76-100. DOI: 10.1006/jmla.1996.0005

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*(4), 803-814. DOI: 10.1037/0278-7393.21.4.803

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. New York: Houghton Mifflin Harcourt.

Scoboria, A., Jackson, D., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General*, *143*(3), 1242-1258. DOI: 10.1037/a0034110

Scoboria, A., Mazzoni, G., & Jarry, J. (2008). Suggesting childhood food illness results in reduced eating behavior. *Acta Psychologica, 128*, 304-309. DOI: 10.1016/j.actpsy.2008.03.002

Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology*, *18*(7), 791-807. DOI: 10.1002/acp.1062

Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., & Hyman, I. E. (2017). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory*, *25*(2), 146-163. DOI: 10.1080/09658211.2016.1260747

Spearman C (1904) The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72–101. DOI: 10.2307/1422689

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*(3), 494-500. DOI: 10.3758/BF03211543

Wade, K. A., Garry, M., & Pezdek, K. (in press). De-constructing rich false memories of crime: Commentary on Shaw and Porter (2015). *Psychological Science*.

Wade, K. A., Garry, M., Read, J.D., & Lindsay, D. S. (2002). A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin & Review, 9*(3), 597-603. DOI:10.3758/BF03196318

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures? *Journal of Applied Psychology*, *82*(2), 247-252. DOI: 10.1037/0021-9010.82.2.247

Watkins, M. J., & Peynircioglu, Z. F. (1990). The revelation effect: when disguising test items induces recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1012-1020. DOI: 10.1037/0278-7393.16.6.1012

Weldon, Corbin, Garavito, & Reyna (2016). Gist is sophisticated yet simple: Fuzzy-trace theory's developmental approach to individual differences in judgement and decision-making. In M. E. Toplak & J. Weller (Eds.). *Individual differences in judgement and decision-making: A developmental perspective*. (pp. 71). New York: Routledge.

Wilkinson, C., & Hyman, I. E. (1998). Individual differences related to two types of memory errors: Word lists may not generalize to autobiographical memory. *Applied Cognitive Psychology*, *12*(1), 29-46. DOI: 10.1002/(SICI)1099-0720(199812)12:7<S29::AID-ACP596>3.0.CO;2-R

Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: Forced confabulation and confirmatory feedback increase false memories. *Psychological Science*, *12*(6), 473-477. DOI: 10.1111/1467-9280.00388

Zhu, B., Chen, C., Loftus, E. F., Lin, C., & Dong, Q. (2013). The relationship between DRM and misinformation false memories. *Memory & Cognition*, *41*(6), 832-838. DOI: 10.3758/s13421-013-0300-2

Zimmerman, M., Ruggero, C. J., Chelminski, I., Young, D., Posternak, M. A., Friedman, M., ... & Attiullah, N. (2006). Developing brief scales for use in clinical practice: the reliability and validity of single-item self-report measures of depression symptom severity, psychosocial impairment due to depression, and quality of life. *The Journal of Clinical Psychiatry*, *67*(10), 1536-1541. DOI: 10.4088/JCP.v67n1007

Table 1. Studies correlating false event suggestion, misinformation, and DRM effects.

Study	**N**	**False Memory Measures**	**Result**
Qin, Ogle, & Goodman (2008, Exp. 2)	86	FES; DRM (12 15-word lists; recall, recognition)	*Null*: statistics not reported
Otgaar & Candel (2011, Exp. 2)*	40	FES; DRM (10 10-word lists; recall, recognition)	*Null*: Believers versus non-believers of false event suggestion on: DRM critical-lure recall $t(38) = -0.88$, $p = 0.38$, $d = 0.29$; DRM critical-lure recognition $t(36) = -0.13$, $p = 0.90$, $d = 0.04$
Otgaar, Verschuere, Meijer, & van Oorsouw (2012)	45	FES; DRM (10 10-word lists; recall)	*Positive relationship*: $F(1,27) = 4.33$, $p < 0.05$, $\eta^2 = 0.14$
Zhu, Chen, Loftus, Lin, and Dong (2013)	432	Misinformation (two events; 12 critical items, recognition); DRM (10 12-word lists; recognition)	*Positive correlation*: $r = .12$, $p = .02$
Ost et al. (2013)	120	Misinformation (one event; three critical items; recall, recognition); DRM (5 15-word lists; recall, recognition)	*Null*: Overall misinformation and: DRM critical-lure recall $r = -.03$, $p = .74$; critical-lure recognition $r = -.03$, $p = .74$
Calvillo & Parong (2016)	160	Misinformation (two events; 12 critical items; recognition); DRM (12 15-word lists; recognition)	*Null*: $r = -.07$, $p = .38$
Monds, Patterson, & Kemp (2016)	87	Misinformation (one event; eight critical items; recall, recognition); DRM (12 10-word lists; recall, recognition)	*Null*: Misinformation recognition and DRM recognition Pearson correlation $p > .14$.
Patihis, Frenda, & Loftus (in press, Exp. 2)	163	Misinformation (two events; 12 critical items; recognition); DRM (20 15-word lists; recognition)	*Null*: $r = .01$, $p = .86$

Exp. = Experiment; FES = False event suggestion; DRM = Deese-Roediger McDermott. * Child participants (*M* age = 9 yrs)

Table 2. Meta-analysis of studies correlating misinformation and DRM effects.

Study	Year	N	Misinformation	DRM	Correlation
Zhu et al.	2013	432	2 events, 12 critical items	10 12-word lists	0.12*
Ost et al.	2013	120	1 event, 6 critical items	5 15-word lists	-0.03
Calvillo & Parong	2016	160	2 events, 6 critical items	12 15-word lists	-0.07
Monds et al.	2016	68	1 event, 8 critical items	12 10-word lists	0.17
Patihis et al.; Exp 2	in press	163	2 events, 12 critical items	20 15-word lists	0.01
Combined	---	943	---	---	0.05

*p < .05

Table 3. Study procedure

	Procedure		Measures administered
Session 1	False food-event suggestion	Pre-manipulation measures	FP, LEI, ABMQ, FHI
	Individual differences		DS; VVIQ; DES-II; Persuadability; AM; GCS; CEQ; MPQ-II; Demographics
Session 2	Misinformation	Present scenes	
	False food-event suggestion	Feedback manipulation	
	False food-event suggestion	Post-manipulation measures	FHI, FP, HB, BB
	Misinformation	Present misinformation	
	False food-event suggestion	Post-manipulation measures	MvB, ABMQ
	Misinformation	Post-manipulation test	Memory test
Session 3	DRM	Present 30 lists	Recall
	DRM	Recognition	Recognition test

Notes: FP – Food Preferences; LEI – Life Events Inventory; ABMQ – Autobiographical Belief and Memory Questionnaire; FHI – Food History Inventory; DS – Disgust Sensitivity; VVIQ – Vividness of Visual Imagery Questionnaire; DES – Dissociative Experiences Scale; AM – Attitudes About Memory; GCS – Gudjonsson Compliance Scale; CEQ – Creative Experiences Questionnaire; MPQ – Multidimensional Personality Questionnaire; HB – Hindsight Bias; BB – Breakfast Behavior; MvB – Memory or Belief

Table 4. Correlations between false food-event suggestion and misinformation effects

	False food-event suggestion					
	Change belief peach yogurt event	Change Sick Dairy	Behavioral intention, peach yogurt	Change preference yogurt factor	Change preference peach factor	Belief [1]
Misinformation						
Correct	-.08	-.07	**-.16**	-.03	**-.17**	-.01
	[-.19,.05]	[-.19,.05]	**[-.28,-.02]**	[-.10,.18]	**[-.29,-.03]**	[-.14,.10]
Mislead	**.15**	.07	**.14**	.02	**.15**	.06
	[.01,.28]	[-.05,.19]	**[.00,.27]**	[-.18,.13]	**[.03,.26]**	[-.07,.19]

N = 237. 1 – Spearman's rho. 95% confidence intervals appear below correlations and were generated by bootstrapping 1000 samples. Correlations for which the confidence interval does not overlap with zero are bolded.

Table 5. Correlations between false food-event suggestion and DRM effects

		False food-event suggestion					
		Change Belief peach yogurt event	Change Sick Dairy	Behavioral intention, peach yogurt	Change preference yogurt factor	Change preference peach factor	Belief [1]
DRM Recall	Correct	-.09	.01	-.05	**.15**	.05	-.08
		[-.21,.04]	[-.12,.15]	[-.17,.08]	**[.03,.27]**	[-.07,.15]	[-.20,.04]
	Critical	.03	-.04	.04	-.01	-.02	.09
		[-.10,.16]	[-.17,.09]	[-.07,.15]	[-.12,.10]	[-.13,.09]	[-.03,.22]
	Intrusions	.03	-.01	.06	.01	-.03	.07
		[-.14,.15]	[-.14,.13]	[-.06,.19]	[-.11,.14]	[-.14,.09]	[-.05,.20]
DRM Recognition	Correct	.00	.01	.07	-.04	.03	.06
		[-.12,.12]	[-.12,.14]	[-.07,.19]	[-.17,.09]	[-.10,.15]	[-.07,.19]
	Critical	.04	.03	.10	-.02	.03	**.23**
		[-.09,.16]	[-.10,.15]	[-.03,.22]	[-.13,.09]	[-.09,.14]	**[.12,.34]**
	Intrusions	.06	.07	**.16**	-.05	.05	.06
		[-.05,.16]	[-.05,.18]	**[.03,.27]**	[-.19,.10]	[-.05,.15]	[-.06,.19]

N = 236. [1] – Spearman's rho. 95% confidence intervals generated by bootstrapping 1000 samples.

Correlations for which the confidence interval does not overlap with zero are bolded.

Table 6. Correlations between misinformation and DRM effects

	DRM					
	Recall			Recognition		
	Correct	Critical	Intrusions	Correct	Critical	Intrusions
Misinformation						
Correct	**.28**	.03	**-.21**	.10	**-.09**	**-.24**
	[.18,.38]	[-.06,.13]	**[-.31,-.11]**	[.00,.20]	**[-.18,.01]**	**[-.33,-.14]**
Mislead	**-.23**	-.04	**.17**	-.05	**.13**	**.21**
	[-.32,-.14]	[-.14.06]	**[.06,.28]**	[-.16,.04]	**[.02,.22]**	**[.09,.33]**

N = 360 for DRM Recall, N = 353 for DRM Recognition. 95% confidence intervals generated by bootstrapping 1000 samples. Correlations for which the confidence interval does not overlap with zero are bolded.

Table 7. Correlations among false memory techniques and individual difference measures

			Dissociative Experiences Scale						
		VVIQ	Total	Amnesia	Depers	Absorb	Persuade	Compliance	CEQ
Food suggestion	Belief peach yogurt	-0.03	**0.17**	0.11	**0.13**	0.06	-0.04	-0.08	0.05
	Change sick dairy	-0.09	-0.04	-0.10	-0.03	-0.11	-0.09	-0.05	0.01
	Party peach yogurt	-0.01	-0.01	-0.03	-0.04	0.06	0.07	**0.13**	-0.08
	Change preference yogurt	0.06	-0.04	0.03	0.02	0.02	0.00	-0.01	-0.08
	Change preference peach	0.07	-0.04	-0.03	-0.11	0.03	0.02	0.06	**-0.10**
Misinformation	Correct	-0.04	**-0.21**	**-0.21**	**-0.13**	0.04	0.06	-0.03	**-0.11**
	Mislead	0.04	**0.19**	**0.21**	**0.12**	-0.04	-0.04	0.09	0.08
DRM recall	Total correct	**-0.12**	**-0.21**	**-0.15**	**-0.18**	0.05	0.09	-0.04	-0.01
	Critical lure	0.03	0.03	0.01	0.04	0.01	0.02	-0.01	0.05
	Intrusions	0.06	**0.18**	**0.13**	**0.15**	0.02	0.03	0.05	**0.11**
DRM recognition	Total correct	-0.05	-0.07	-0.06	-0.04	-0.03	**0.12**	0.05	0.07
	Critical lure	0.04	**0.18**	**0.18**	**0.18**	0.02	0.07	0.05	**0.16**
	Intrusions	0.04	**0.18**	**0.17**	**0.14**	-.04	0.04	0.02	0.06

Correlations with p < .05 are indicated in bold font. VVIQ = Vividness of Visual Imagery Questionnaire. Depers = Depersonalization. Absorb = Absorption. Persuade = Persuadability. CEQ = Creative Experiences Questionnaire.

Appendix 1. False food-event suggestion, misinformation and DRM Internal consistency reliability

		Items	Alpha
False food-event suggestion	Food preference Time 1		
	Yogurt factor	4	.87
	Peach factor	3	.79
	Fish factor	4	.83
	Peanut factor	2	.82
	Belief Peach yogurt, Time 1		
	Peach factor	2	.71
	Pickle factor	2	.63
Misinformation	Correct	36	.60
	Mislead	24	.64
	Foil	36	.28
DRM Recall	Correct	450	.94
	Critical lure	30	.76
DRM Recognition	Total correct	30	.83
	Critical lure	30	.85
	Intrusions	30	.73

1 – Spearman-Brown.

Appendix 2. False food-event suggestion Test-retest reliability, Pearson correlations

(Control group, scores should not be impacted by intervening manipulation)

		Items	Control (N = 90)	Suggestion (N = 286)
Food history	Sick on dairy	1	.58	.46 [.35,.56]
Food preference	Yogurt factor	4	.88	.82 [.77,.86]
	Peach factor	3	.82	.70 [.62.77]
	Fish factor	4	.95	.94 [.92,.96]
	Peanut factor	2	.78	.80 [.74,.85]
Belief	Peach yogurt	2	.38	.39 [.28,.50]
	Pickle	2	.45	.38 [.26,.50]