A Mega-analysis of Memory Reports from Eight Peer-reviewed False Memory Implantation Studies

Alan Scoboria (University of Windsor)

Kimberley A. Wade (University of Warwick)

D. Stephen Lindsay (University of Victoria)

Tanjeem Azad (Kent State University)

Deryn Strange (John Jay College of Criminal Justice)

James Ost (University of Portsmouth)

Ira E. Hyman, Jr. (Western Washington University)

In Press, *Memory*, November 9, 2016. This is not the official copy of record; the final published version may differ slightly.

Author note: Please direct correspondence to A. Scoboria, Department of Psychology, 401 Sunset, Windsor, Ontario, Canada, N9B 3P4, scoboria@uwindsor.ca; K. Wade, Department of Psychology, University of Warwick, University Road, Coventry, CV4 7AL, United Kingdom, K.A.Wade@warwick.ac.uk; S. Lindsay, Department of Psychology, University of Victoria, 3800 Finnerty Road, Victoria BC, Canada,V8W 2Y2, slindsay@uvic.ca. This research was supported by a Natural Science and Engineering Research Council of Canada Discovery Grant to the first author (RGPIN/327570−2012). The authors thank the many research assistants who contributed to this research.

Abstract

Understanding that suggestive practices can promote false beliefs and false memories for childhood events is important in many settings (e.g., psychotherapeutic, medical, legal). The generalizability of findings from memory implantation studies has been questioned due to variability in estimates across studies. Such variability is partly due to false memories having been operationalized differently across studies and to differences in memory induction techniques. We explored ways of defining false memory based on memory science and developed a reliable coding system that we applied to reports from eight published implantation studies (N=423). Independent raters coded transcripts using seven criteria: accepting the suggestion, elaboration beyond the suggestion, imagery, coherence, emotion, memory statements, and not rejecting the suggestion. Using this scheme, 30.4% of cases were classified as false memories and another 23% were classified as having accepted the event to some degree. When the suggestion included self-relevant information, an imagination procedure, and was not accompanied by a photo depicting the event, the memory formation rate was 46.1%. Our research demonstrates a useful procedure for systematically combining data that are not amenable to meta-analysis, and provides the most valid estimate of false memory formation and associated moderating factors within the implantation literature to date.

Keywords: false memory, suggestion, mega-analysis

A Mega-analysis of Memory Reports from Eight Peer-reviewed False Memory Implantation Studies

Memories are precious. They bond relationships, contribute to a sense of identity, and shape current decisions and future planning (Bluck, Alea, Habermas & Rubin, 2005; Brown, Schweickart, & Svob, 2016; Conway & Loveday, 2015; Pillemer, 1998). Memories may also seem eternal, like cherished photographs in an album we peruse from time to time. Our memories play major roles in making us who we are. Our beliefs about our personal histories both reflect and constitute central aspects of ourselves. Practitioners (e.g., police officers, medical personnel, career guidance counsellors, historians, political scientists), in a variety of everyday settings routinely rely on individuals' autobiographical memory reports sometimes basing extremely consequential decisions on individuals' reports of their personal histories.

Yet remembering the past is a complex phenomenon that is subject to error (Schacter, 2013). The malleable nature of human memory has led some researchers to argue that our memory systems are not oriented toward flawlessly preserving our past experiences. Indeed, many researchers now agree that remembering is, to some degree, reconstructive (Brewer, 1986). Current theories propose that our capacity to flexibly recombine remembered information from multiple sources—such as distributed memory records, inferences, and expectations—helps us to solve current problems and anticipate future events (Johnson & Sherman, 1990; Newman & Lindsay, 2009; Suddendorf & Corballis, 1997; Szpunar, Addis, McClelland, & Schacter, 2010). One implication of having a reconstructive and flexible memory system is that people can develop rich and coherent autobiographical memories of entire events that never happened (Bernstein & Loftus, 2009).

The existence of illusory autobiographical memories and false beliefs about the personal past has profound implications for psychology and for other disciplines. Thus it is crucial for

psychological scientists to understand the mechanisms that underlie the development of mistaken beliefs and illusory recollections and to address critical questions. For example, are there reliable ways to differentiate between accurate and inaccurate autobiographical reports? If so, are there ways to correct mistaken memories or to erase or at least disempower illusory memories?

In this article, we revisit questions about the conditions under which participants in studies of false autobiographical memory come to believe in and remember fictitious childhood experiences. The study of false memories for autobiographical events in adults originally stemmed from concerns that suggestive therapeutic practices might foster the development of false childhood memories, particularly memories of abuse (Lindsay & Read, 1994; Loftus, 1993; Patihis, Ho, Tingen, Lilienfeld, & Loftus, 2014). We now know that suggestive techniques can indeed lead people to develop detailed false autobiographical memories (Loftus & Bernstein, 2005). Such false memories can affect people in a variety of ways, leading to changes in views of the self and views of others, and to changes in social behavior such as pursuing legal charges against a parent. Yet some memory-trauma therapists continue to use ‘memory recovery’ techniques that may put clients and their families at risk (Cara, 2014). Moreover, recent studies of mainstream clinicians and lay therapists show that many still hold inaccurate views about how memory works (Ost, Easton, Hope, French, & Wright, in press; Patihis et al., 2014).

The most powerful procedure that researchers use to study the development of false childhood memories is variously called the ‘lost-in-the-mall’ (Loftus & Pickrell, 1995), ‘familial-informant false narrative’ (Lindsay, Hagen, Read, Wade, & Garry, 2004), or ‘memory implantation’ (Wade, Garry, Read, & Lindsay, 2002) methodology. These studies typically merge suggestive techniques with social pressure to lead participants to report believing and remembering that a suggested pseudoevent had actually occurred. Briefly, in studies of this type

adult participants are told that the researchers are interested in how people recall childhood events. The participants are then presented with descriptions of a set of childhood events that were ostensibly provided by trustworthy family members (such as parents or siblings). Amongst these events is one false event created by the researchers (familial informants usually verify that to their knowledge the participant never experienced this event in childhood). Over two or three sessions – typically over the course of a week – participants are encouraged to recall the childhood events using various memory recovery techniques employed in trauma-memory-oriented therapy (e.g., guided visualization). Researchers then assess the extent to which participants form false memories by the final interview session. This method remains the most rigorous experimental procedure for studying the creation of false autobiographical memories, and in this paper we use the term *false memory* to refer to the outcomes of this specific methodology, unless otherwise noted. We use the term *false belief* to identify the development of a belief that the suggested event occurred in the past, whether or not the person experiences a vivid memory of the event.

The primary goal of memory implantation studies has been to establish which suggestive procedures enhance or inhibit false memory formation. To answer this question, researchers must determine when a false memory is present and estimate the frequency with which people form false memories under different conditions. To this end, researchers have proposed operational definitions for deciding when false memories have or have not developed. Independent judges typically read transcripts of participants' memory reports and determine whether they have developed false memories according to a set of guidelines specific to the study. If judges disagree on a categorization they may categorise the response in the more conservative category or resolve the disagreement via discussion. Many studies also ask participants to rate the strength

of their memory and other variables related to the subjective experience of remembering (such as the extent to which they mentally relive the event as they bring it to mind, with varying degrees of sensory/perceptual, valence, and emotional qualities). Researchers sometimes use these subjective ratings to validate the judges' categorizations.

Although memory implantation studies have used these general procedures, the absolute rate of false-memory creation has varied dramatically across studies and across conditions. Among published studies involving healthy young adults, the rate of false recall has ranged from 0%[1] (Pezdek, Finger & Hodge, 1997) to 65% (Lindsay, Hagen, Read, Wade, & Garry, 2004). It is likely that several different factors contribute to differences in the rate of false memories across studies. For example, the nature of the suggested event and the persuasiveness of the research assistant probably play roles. Another likely contributor, and a primary focus of this paper, is variation across studies in how false memories were operationalized.

In the study we report here, we developed and implemented a new coding scheme for determining whether or not participants reported false memories. By using a single measurement tool to assess over 400 memory-report transcripts from eight published studies, we developed an improved estimate of the frequency with which false memories occurred. We also gained insight into the relative power of different suggestive techniques for creating false memories. Both of these outcomes will improve our understanding of the suggestive procedures and the mechanisms that underlie illusory memories and false beliefs.

[1] The 0% memory formation reported by Pezdek, Finger & Hodge (1997) resulted when individuals were intentionally presented with a low plausibility event (receiving an enema) and interviewed by a sibling rather than a trained researcher; see Scoboria, Mazzoni, Kirsch & Relyea (2004), for personal plausibility ratings for 'enema' and other events. Low plausibility is associated with inhibition, but not elimination, of false belief and memory formation (Scoboria, Mazzoni, Jarry & Shapero, 2012). But unremembered childhood abuse is viewed as personally plausible by a sizeable minority of American undergraduates (Pezdek & Blandon-Gitlin, 2008), and the plausibility of initially low plausibility events can be increased (Mazzoni, Loftus, & Kirsch, 2001).

Approach to Combining the Studies

At first glance, it may appear that using meta-analytic procedures would be a valid approach to arrive at combined estimates of memory formation across studies. Indeed, meta-analysis often affords powerful explorations of relationships between variables. But when combining studies it is critical that the dependent variables were measured in a sufficiently similar manner so that they can either be combined directly or placed onto standardized scales. For a variety of reasons this cannot be accomplished with memory implantation studies.

The coding systems used to determine the prevalence and strength of false memories have varied widely across memory implantation studies. Table 1 provides a summary of the various ways in which researchers have operationalized false memory. It is clear from the table that the operationalizations draw on different dimensions. For example, some researchers only required that participants state that they remember the event, other researchers required that individuals provide details that extend beyond the suggested material, and still other researchers required that participants make a rating above a criterion on a self-report scale, to classify reports as false memories. Studies that have common dimensions often vary in the degree to which each dimension is emphasized. Furthermore, the highest category, often called ‘full’ or ‘clear’ false memories, is not defined consistently. The problem is compounded by the differential use of intermediate categories, such as ‘images without memories’ or ‘partial false memories’.

These concerns are present even before we consider that multiple independent judges in different locations were tasked with categorising participants’ false memory reports. Judges do not always agree, for instance, on how to classify participants’ reports, and the exact procedures used for handling disagreements may lead to more or less conservative criteria being applied. Thus even when different labs use similar coding schemes, it is difficult to tell whether the

judges applied the coding scheme consistently. In short, false memories in memory implantation studies have been defined and measured in a variety of ways, which makes it difficult to determine just how prevalent they are across studies.

Given the variation in coding schemes across studies, meta-analysis is not an appropriate approach for combining data from memory implantation studies. Indeed, there is currently no valid method for combining summary statistics originating from scales that differ in ways that are unknown. An alternative is to re-code the original data using a single validated system and to conduct a 'mega-analysis' of the resulting dataset. Mega-analysis involves the combination of participant level data from multiple studies and hence stands as a meta-analysis of individual data, which Olkin (1995) argued is the highest form of empirical evidence when combining studies. Providing that researchers have used the same dependent measures and are willing to share the raw data, mega-analysis has the advantage of accruing larger samples, thereby enhancing statistical power and permitting examination of research questions that cannot be addressed in individual, smaller studies. Mega-analysis has been increasingly used in areas such as genetics, neuroimaging, and treatment of psychopathology (e.g., de Maat et al., 2008; Hallahan, et al., 2011; see Sternberg et al., 2006, for a study of the effects of family violence on child behavior, in which combining data from multiple studies permitted asking of new questions). Bernstein, Scoboria, and Arnold (2015) published a mega-analysis (8 studies; combined N = 1,369) of a different method that involves providing suggestions about false childhood food experiences. In the present study, we examined transcripts from false memory implantation studies at the participant level of analysis.

Identifying false autobiographical memories

We grounded our definition of false memory in established theories of remembering (e.g., Brewer, 1996; Conway & Pleydell-Pearce, 2000; Jacoby, Kelley, & Dywan, 1989; Johnson, Hashtroudi, & Lindsay, 1993; Ross, 1989; Rubin, 2006; Tulving, 1983). The experience of remembering a past event is the result of multiple cognitive systems and processes. Inputs from numerous sources (e.g., records of prior perceptual processing, of prior thoughts, feelings, general knowledge, beliefs, goals, feelings of familiarity, etc.) give rise to an experience of remembering the past. One key component is *recollection*, which involves the subjective experience of mentally reliving a past event. Mental representations experienced as recollections contain information about perception, place, actions, persons, and so forth. Another key component is *belief in occurrence*: remembering also entails the belief that an event genuinely occurred in one's personal past (Brewer, 1996; Scoboria, Talarico, & Pascal, 2015). It is possible to believe that one experienced a particular event in the past without recollecting it (e.g., you believe that you were born although you don't recollect the event). It is also possible to experience vivid recollections that are not believed to represent genuine past events (Scoboria et al., 2014). Thus definitions of false memory should address both belief about the occurrence of the suggested event and the subjective experience of recollecting the event.

To develop our coding scheme for defining 'false memory', we began with these observations and with themes previously used in the literature. We chose seven concepts: [1] verbal statements of 'remembering'; [2] acceptance of suggested information, which indirectly indicates a degree of belief in the occurrence of the event; [3] elaboration beyond suggested information, which indicates extension of a suggestion beyond acceptance of and compliance with suggested material; [4] presence and quality of mental imagery; [5] coherence of memory narratives; [6] evidence of emotional experience; and [7] no rejection of the suggested event.

Each of these dimensions could, we believed, be reliably assessed from the information available in transcripts of participants' memory reports. We do not claim that these concepts represent a comprehensive list of the dimensions associated with remembering. Rather, we suggest that they adequately reflect current thinking about the characteristics of autobiographical memories, and that they are sufficient to determine the presence or absence of implanted memories.

Statements of remembering. Some studies have used statements such as "I remember" to designate the presence of an implanted memory. Yet this criterion, alone, is insufficient. Just as the statement 'I confess' should not be used to infer that someone genuinely believes that he or she committed a crime (Kassin, 2012), the statement 'I remember' should not be used to infer that someone is indeed experiencing vivid recollection or belief that the event occurred in the past. Issues such as experimental demand and compliance with the research procedure may produce verbal reports that do not reflect peoples' beliefs or recollective experiences about the past.

A related concern is that people vary in the ways in which they talk about their memories, partly due to differences in how they have learned to communicate about the past (Fivush, Habermas, Waters, & Zaman, 2011). Variability in linguistic usage leads to instances in which recollection may be present although the person does not appear to report recollection, or vice versa (see Otgaar et al., 2013). In short, verbal statements of remembering may be a starting point for identifying false memories, but alone are not adequate for determining that an event is remembered.

Acceptance of the suggested event. At the start of false memory studies, participants typically initially indicate that they do not remember the suggested event. When a false event is first presented, some participants reject it outright, saying "No, that never happened to me."

Often such people say something like, “That must have happened to my brother,” as a way to account for the belief that the informant wrongly claimed that the event happened to them. Such participants are unlikely to develop a memory of the false event unless they somehow come to accept that the suggested event occurred (Hyman & Kleinecht, 1999). People will not search for event-consistent thoughts and images or interpret them as memories of an event until they find it at least slightly plausible that the event could have transpired (Mazzoni, Loftus, & Kirsch, 2001; Scoboria, Mazzoni, Jarry, & Shapero, 2012). The first step of the journey toward a false memory is some degree of acceptance that the suggested event could have happened. This may range from minimal acceptance of the possibility that the event may have occurred, to complete belief that the event definitely occurred (Pezdek, Blandon-Gitlin, Lam, Hart & Schooler, 2006; Thomas & Loftus, 2002; see Scoboria et al., 2004, for more on the relationship between plausibility and autobiographical belief). The fact that acceptance is not explicitly stated as a criterion in some studies may reflect the questionable assumption that statements of remembering imply that someone has accepted the event.

Elaboration beyond the suggestion. Many prior definitions of false memory require that participants elaborate beyond the details provided in the suggestion. Participants often embellish their reports with additional, event-relevant information and this elaboration provides some indication that participants are making an effort to remember and that they are not merely complying with experimenter demand. The source monitoring framework posits that people evaluate event-relevant information that comes to mind during attempts to remember to determine the extent to which that information resembles a genuine memory (Johnson et al., 1993; Lindsay, 2008). Thoughts and images that arise from inference or imagination often retain the sense of having been internally generated (what the source monitoring framework terms

markers of effortful cognitive operations), but sometimes internally generated information is misattributed to past experience. Various aspects of the implantation procedure, such as telling participants that a family member provided the events, may encourage participants to lower their criteria for accepting internally generated thoughts and images as genuine memories (Wade & Garry, 2005).

Self-generated event details can also lead to activation of additional relevant material in memory (Greenwald & Banaji, 1989) and to more frequent false memories (Hyman et al., 1995). Conversely, when people fail to retrieve sufficient information or retrieve event-inconsistent details, they may reject the event (Dodson & Schacter, 2001; Ghetti & Alexander, 2004; Lampinen & Odegaard, 2006). Hence elaboration beyond the details of the suggestion is an important facet of memory formation.

Imagery. In many implantation studies participants are encouraged to create perceptual imagery for the false event. Imagery has long been associated with remembering (Brewer, 1996; Greenberg & Knowlton, 2014; Johnson, Foley, Suengas, & Raye, 1988), and the presence and strength of imagery is an important dimension to consider when examining memory formation. Theories of memory formation that build upon the source monitoring framework (Johnson et al., 1993) posit that some suggestive practices foster the creation of imagery which may subsequently be misattributed to memory (Hyman & Kleinknecht, 1999; Lindsay & Read, 1994). Recent findings support the view that the mental simulation of events recruits overlapping neural networks, regardless of whether those events are attributed to memory, to imagination, or to some other source (Addis, Pan, Vu, Laiser & Schacter, 2009; Schacter, Camberlain, Gaesser, & Gerlach, 2012). However the presence of vivid imagery alone is not indicative of recollection;

the imagery must also be attributed to prior experience (erroneously in the case of false memories).

Narrative coherence. Rubin (2005) proposed that narrative is a key property of remembering that serves to organize autobiographical memory. Memories are frequently expressed in a narrative form, with a beginning and end, characters, settings, events, and so forth. Narrative is also the means by which memories for events are typically communicated. Thus the degrees to which memory reports appear to be complete, coherent, and thematically consistent are important dimensions of both the subjective and social aspects remembering (Brewer, 1986; Pillemer, 1998).

People are also unlikely to accept suggested events as genuine past experiences if the events do not coincide with broader self-narratives. When people attempt to recall a past event, if a memory does not come quickly to mind they may try to identify the broad lifetime period and context that likely surrounded the event (Brown et al., 2016; Conway & Pleydell-Pearce, 2001; Neisser, 1987). These more general details may then cue recollection for a specific episode. The more inconsistent an event is with one's current self-views, however, the less likely a person is to accept thoughts, images and feelings that come to mind as evidence of remembering (cf. Ross, 1989). Suggested events must make sufficient sense in relation to a person's current perceptions of their past. When people view an event as being consistent with their personal history, their internal representation of the event is more likely to feel coherent. Moreover, when they talk about the event, others will judge it to be coherent as well. Thus narrative coherence should be a marker of the extent to which people accept, integrate, and elaborate on the suggested material to produce an integrated memory narrative (Neisser, 1994).

Emotion. Given that people experience emotional states when experiencing events, it is not surprising that people also recall past emotional states and experience emotional reactions when recalling events (Christianson & Safer, 1996). Based on findings that emotional intensity at retrieval is associated with strength of autobiographical memories (Talarico, LeBar, & Rubin, 2004), we opted to code intensity of emotion in reports as an additional indicator of the development of false memories.

No rejection of the suggested event. Participants in memory implantation studies are encouraged to construct vivid imagery while narrating aloud in order to retrieve unremembered events. In some cases individuals may provide an otherwise compelling report, but also make statements that indicate that they reject having any memory for the suggested event. It is obvious that such cases should not be categorized as false memories.

Analytic approach

To conduct the mega-analysis we solicited transcripts from published peer-reviewed false memory implantation studies. We developed and validated a system for coding false event narratives, and then applied the system to the transcripts from the studies. This approach permitted us to examine four novel outcomes. First, the combined dataset allowed us to take a data-driven approach to defining the presence of false memories. Second, we were able to combine memory formation rates across studies in a more valid manner. Third, the statistical power afforded by the larger dataset enabled us to examine the convergence between coded false memories and participants' self-report ratings. Finally, we were able to explore what study characteristics were associated with false memory formation.

Method

Development of the coding system

Using the themes described above, we developed and validated a coding system for deciding when participants had experienced features associated with remembering. We created seven primary items (statements of remembering, acceptance, elaboration, imagery, coherence, emotion, and no rejection of the suggested event). For each dimension, raters used a 4-point rating scale, ranging from 'not present' to 'highly present'. For three items (imagery, emotion, elaboration), raters first indicated whether the phenomenon was present or not, followed by 4-point scaled questions about the amount and quality of the feature (imagery amount, imagery quality, emotion amount, emotion intensity, elaboration amount). We created an item to assess whether a memory was present based on the entire transcript (judged memory; rated no, yes/partial, yes/complete). The no rejection of suggested event item was rated as yes (rejected) or no (not rejected). We worded all items with an eye towards conceptual clarity and created an accompanying set of definitions for each item.

Two assistants from the first author's lab were trained in the use of the system. The raters worked with the authors to code event transcripts of true and false events from unpublished studies. We completed three cycles of coding and revising item wording until the raters provided consistent judgments. The same raters then independently coded a new set of 68 unpublished false memory transcripts (taken from pilot work and unfinished studies). We calculated the percentage of ratings that were identical between raters and examined frequencies for ratings that differed by two or more points on the scales, because these represent instances in which the raters read the transcript quite differently. We retained items with agreement above 70% and for which discrepancies between item ratings were less than two scale points in at least 90% of cases. See Table 2 for the final items.

Obtaining and coding false memory transcripts

We approached researchers in Canada, New Zealand, the United States of America, and the United Kingdom who had published peer-reviewed memory implantation studies involving adult participants. We asked them to provide the anonymized, final-session transcripts for the false events from their published studies. The transcripts no longer existed for four published studies so we were unable to include data from these studies in our analyses (see the discussion below for further information about these studies). Ultimately we obtained transcripts from eight published studies involving a total of 423 participants: Desjardins and Scoboria (2007; *N* = 43; suggested event, trouble with teacher); Garry and Wade (2005; *N* = 44; hot air balloon ride); Hessen Kayfitz and Scoboria (2011, *N* = 81; hot air balloon ride); Hyman and Pentland (1996; *N* = 64; spilled punch bowl on parents of the bride at a wedding); Lindsay et al. (2004; *N* = 45; trouble with teacher); Ost et al. (2005; *N* = 26; one of eight events); Strange et al. (2008; *N* = 100; hot air balloon ride); Wade et al. (2002; *N* = 20; hot air balloon ride). The distribution of gender across the total sample was 71.2% female, 25.9% male, 2.8% gender not recorded. Age was recorded for 65% of the sample; *M* = 20.7, range 17 to 42, *SD* = 2.9 (age was not available for the Hyman, Lindsay, or Ost studies but they too used university undergraduates as participants). We compared samples on the available demographic characteristics and found no statistically meaningful differences.

To ensure that each item in the coding scheme was rated independently, we (AS, KW, SL/TA) trained separate pairs of raters to code each item. For example, one pair in AS's research lab rated Elaboration, and another pair in KW's lab rated Coherence, and so on. Pairs were trained in the use of an item and coded practice events using unpublished false memory transcripts until they achieved consistency of 75% or greater. Each pair then independently rated the 423 false memory transcripts. Initial inter-rater agreement was good: Elaboration – presence

93.6%, amount 88.2%, relation 92.8%; Acceptance – 84.4%; Imagery – presence 90.7%, amount 87.8%, quality 82.0%; Emotion – presence 94.2%, amount 72.0%, intensity 100%; Coherence – 76.7%; Judge remembered – 79.7%; Statements of remembering, 100%; No rejection, 100%. Inconsistent ratings were resolved via discussion and in rare cases of continued disagreement the more conservative rating was selected.

Defining 'False Memory'

We explored criteria for defining when false memories were present. We began with the following two assumptions as discussed above. First, participants must accept the suggestion to some degree. Second, as participants move from saying 'I do not remember that' to 'Now I remember…', they must report additional imagery or otherwise elaborate beyond the suggested material. Of course, the extent to which these characteristics need to be present to define a participant's report as a memory remains an open question.

Correlations and 95% confidence intervals between judges' acceptance, elaboration, imagery, and judged memory ratings are provided in Table 3. Acceptance showed the highest relationship with judgments of memory, consistent with the idea that a degree of acceptance of the suggestion is a necessary dimension in memory formation. The high correlation (.73) between ratings of imagery and elaboration suggests that these concepts reflect highly overlapping aspects of the reports. In light of these relationships and the preceding discussion, we developed definitions of "false autobiographical memory" that focussed on the extent to which participants accepted the suggested event.

Next we examined various combinations of the acceptance, elaboration, and imagery variables, and the associated mean ratings on other judged variables (emotion engagement, coherence, and judged memory) for each combination. Table 4 displays the cumulative

frequencies for each combination. The first row includes cases for which acceptance of the suggestion was coded as zero. The ratings for all variables in this group were extremely low. The second row includes cases in which any level of acceptance was present, and subsequent rows incorporate higher levels of acceptance and introduce different levels of elaboration and imagery.

The item means indicated a difference between *any acceptance* and *moderate acceptance* of the event. Based on this observation and the idea that memories typically include a degree of acceptance of the event plus a memory-like image and elaboration beyond the suggested material, we placed the minimal criterion for presence of a memory as *moderate acceptance* plus any level of both elaboration and imagery (Table 4, row 4). For the cases classified as memories, increasing the thresholds for elaboration and imagery from *any* to *moderate* resulted in notable increases in scores. Thus we used this threshold to differentiate between two groups of false memories. The first group was named "partial false memory" (moderate acceptance with any elaboration and any imagery) to reflect strong acceptance but lesser presence of other key features associated with memory. The second group was named "full false memory" (moderate acceptance with moderate elaboration and moderate imagery, row 5), to indicate that all key features were present. We defined those cases judged with a high level of acceptance of the suggestion and moderate imagery and elaboration as "robust false memory" (row 6). In previous studies, false memories were classified as either partial or full. Combining the full and robust categories would be similar to the full memories from earlier studies. One advantage of this mega-analysis is the finding that false memories represent a continuum in terms of acceptance and elaboration. With the mega-analysis, we were able to differentiate that even within the group of false memories typically classified as full memories, there are many that display a much greater degree of acceptance and elaboration. Regardless of whether the event met the criteria for

partial, full, or robust false memory, if the participant stated at the end of the interview that s/he did not have a memory of the event, we coded the event as 'rejected.' This resulted in the following categories: rejected, no memory, accepted (to some degree), partial false memory, full false memory, and robust false memory.

Results and Discussion

False memory formation across studies

When we categorized the 423 memory reports according to the scheme described above (see Figure 1), 11.1% of the sample met the criteria for Robust memory, 10.8% met the criteria for Full memory, and 8.5% met the criteria for partial false memory, resulting in a combined false memory rate of 30.5%. Of the remainder, 23.0% were judged as having accepted the suggestion as true to some degree but did not meet the criterion for remembering. Another 10.4% were judged as having strong mental representations but the participant explicitly rejected having a memory. For 36.2% there was no evidence of acceptance of the suggestion or memory formation. Thus, approximately one-third of participants showed evidence of a false memory, and more than half showed evidence of believing that the event occurred in the past.

Statements of remembering the suggested event

As we noted previously, relying on participants' verbal utterances of remembering (e.g., "I remember being lost…") to determine remembering is problematic. For this reason, we did not include statements of remembering in our definition of false memories. We did, however, calculate the percentage of participants in each memory category who made clear statements of remembering. In the Full and Robust categories, 91.2% (n = 85) of cases included a statement of remembering. Of the eight cases that did not include such a statement, in six cases the participant spoke in ways that appeared to reflect remembering but did not use explicit language indicating

such. For the remaining two cases, the participants shifted into describing the event in the historical present at some point in the narrative (e.g., “I bump into the table with my shoulder, and the bowl moves…”), which has been associated with the experience of remembering (Pillemer, 2003).

In the Partial memory category, 36.1% (n = 13) of participants made statements about remembering. Some participants spoke vaguely, which may have made their transcripts more difficult to classify. Other participants explicitly expressed an incomplete sense of remembering, and often indicated that they were retrieving fragments of memories or were vividly seeing isolated details in their mind. As such, they went beyond speculation to express a partial sense of remembering. Thus some of these cases may reflect full memories, but due to their ambiguity they were appropriately coded into the more conservative category.

For the Acceptance category, 8.2% (n = 8) of cases made statements of remembering. Of these, two participants were uncertain if the memory originated in the suggested event or another experience; another two could be interpreted as either indicating imagining or remembering; and for four cases the transcripts presented a mixture of both memory claims and denial of memory. In the No Memory category, no cases were rated as providing statements of remembering.

There are three points to note here. First, participants’ statements of remembering corresponded well, but not perfectly, with our criteria for defining false memories. Participants categorized as having full or robust false memories almost always made statements about “remembering” the event. Second, our criteria appropriately and conservatively allocated participants with vague or fragmented memory representations into the partial memory group. Finally, a notable minority of participants did not make explicit statements of remembering but still may have remembered the suggested event.

Relationship of additional judged items to false memory status

Judges rated the characteristics of the 423 reports on six additional items that were not used to determine the false memory categories: emotion amount, emotion intensity, coherence, elaboration relationship (how closely the additional information was to the original event), imagery quality, and judged memory (whether the rater thought that the participant 'felt like' they had a memory). Table 5 presents descriptive statistics for these items, by the false memory categories defined above. We use 95% confidence intervals to discuss differences on these items between the memory categories (Cumming, 2013). We focus on the 'No-memory' to 'Robust' memory categories first, then examine the 'Rejected' category separately below. We found a strong convergence between the memory categories and the judges' ratings on all six items. For example, reports that were categorized as Robust false memories scored higher on the six items than reports categorized as Full. The rank order of the categories was consistent across all items, with means going from lowest to highest value from the No memory to the Robust category; although some items did not significantly differ across categories. Contrasting the contiguous groups, the Full and Robust categories were differentiated only by the coherence and judged memory items. The Partial and Full categories were differentiated by the coherence, imagery quality, and judged memory categories. The Accept and Partial categories were differentiated by coherence, elaboration relationship, imagery quality, and judged memory. Finally, the No memory and Accept categories were differentiated by all items except emotion engagement.

To summarize, mean ratings for all items increased numerically across the No-memory to Robust categories, respectively. Coherence ratings differentiated all categories, and Imagery Quality and Judged Memory ratings differentiated a number of categories.

Rejected category. Participants in the Rejected memory category (10.4% of the sample) were coded as having vivid mental representation for the suggested event, but ultimately rejected remembering the event. Not surprisingly, Judged Memory ratings for these participants were invariably zero. This group's ratings for elaboration relationship resembled those who were deemed to have Full or Robust memories, which is consistent with the argument that these reports look like vivid memories to an observer. Emotion and imagery quality in the Reject category resembled the Partial category, whereas coherence ratings for the Rejected category were lower than those for the Partial category. The pattern of ratings in the rejected category is interesting in that they highlight the difficulty that judges face in 1) determining when someone is remembering an event, and 2) determining the underlying psychological basis for such memory statements.

Participants' subjective ratings for suggested events

We now turn to participants' responses to self-report items, which can serve as another method for validating judges' ratings. In memory implantation studies, participants are often asked to rate the phenomenological characteristics of their memory reports, such as the extent to which they see the event in their mind or relive the original emotions associated with the event. These ratings enable participants to convey their subjective sense of remembering. The self-report items used most frequently across the eight studies included: visual imagery, auditory imagery, emotional intensity, narrative coherence, remember/know, prior instances of rehearsal, confidence in memory, and belief in the accuracy of memory. We contrasted our memory categories using the self-report items that were completed by at least 50% of the combined sample. Three items met this criterion: [1] Remember/know (75.9%), [2] Rehearsal (75.9%), and

[3] Reliving (58.6%). Figure 2 shows the mean rating for these items according to memory category.

Remember/know and reliving. Remember/know was measured on a 7-point scale, and assessed the degree to which participants' episodic recollection was associated with an event versus knowing that an event happened without associated recollection; recent research has found that this item is an excellent marker of subjective recollective experience (Scoboria et al., 2014). Reliving assesses the extent to which the event representation is associated with an experience of reliving the event, and was also measured on a 7-point scale. Reliving and Remember/know ratings were lowest in the Reject and No Memory groups, were somewhat higher (differing from the scale floor) in the Accept and Partial groups, and were notably higher in the Full and Robust groups. We calculated standardized effect sizes to contrast the No Memory group to the combined Accept/Partial groups (R/K, d = 0.38 [95% CI .11, .65][2]; Reliving, d = 0.57 [.26, .89]), No Memory to the combined Full/Robust groups (R/K, d = 1.29 [.97, 1.61]; Reliving, d = 1.34 [.99, 1.69]), and the combined Accept/Partial groups to the combined Full/Robust groups (R/K, d = 0.92 [.60, 1.24]; Reliving, d = 0.76 [.39, 1.12]). The agreement between the coded ratings and these subjective ratings provides evidence of convergent validity for our coding system.

Rehearsal. The rehearsal item assesses how frequently participants thought or talked about an event in the past on a 7-point scale. Because the suggested event is novel, ratings on this item should be low. This was the case, except that the Robust memory group showed a weak endorsement of the item. Those in the Robust memory group may have given slightly higher ratings due to having thought about the event more during the study.

[2] Confidence intervals on Cohen's d were calculated using the SPSS macro by Wuensch (2012).

Study characteristics and false memory

Within the memory implantation literature over time we see the emergence of increasingly powerful techniques for fostering false memories. We examined the correlation between the judged ratings and factors that varied across conditions within and between the eight studies. Table 6 shows the judged items and how each item correlated (and associated 95% confidence intervals) with three factors that showed statistically meaningful relationships: [1] Providing participants with idiosyncratic self-relevant information (e.g., including details in suggestions such as the name of an actual teacher), [2] Encouraging participants to imagine the suggested event during testing sessions, and [3] Using doctored photographs to depict the suggested event. We also examined whether providing participants with both a narrative and a photo was related to memory rates, and did not find any meaningful correlations (range - .07 to .05). Whether significant or not these correlations must be interpreted with caution due to the presence of multiple confounds. Nonetheless, they offer some intriguing leads, as detailed below.

Providing self-relevant information. Desjardins and Scoboria (2007) proposed that studies with higher false memory rates provide participants with idiosyncratic self-related information within the suggestion, and they reported a study that yielded support for that hypothesis. We coded whether idiosyncratic information was (6 studies, N = 174[3]) or was not (5 studies, N = 249) provided to participants. The presence of idiosyncratic information correlated positively with all of the judged categories except for emotional experience. The correlations ranged from .16 (elaboration relationship, judged memory) to .30 (imagery amount). Idiosyncratic self-relevant information appears to enhance visual imagery, increase elaboration on suggested material, and foster a cohesive narrative. The false memory rate (Robust, Full, and Partial

[3] Some published papers contribute cases to both categories due to the manipulation of variables within studies.

combined) was 39.6% [95% CI: 32, 47] when self-relevant information was present, and 24.0% [19, 30] when it was not; $prop_{diff}$ = .24 [.07, .24].

Instructing participants to imagine. Many researchers have discussed the influence of imagination and of vivid imagery on memory formation (Hyman & Pentland, 1996; Mazzoni & Memon, 2003). We coded whether participants were (6 studies, N = 265) or were not (3 studies, N = 158) encouraged to systematically imagine the suggested event during the study procedure. Instructions to generate imagery correlated positively with all of the judged variables. Most of the correlations fell between r = .21 and .28, indicating weak but reliable relationships. The correlation with the Coherence item was somewhat higher at r = .38. The false memory rate was 37.3% [95% CI 32, 43] when imagery was present in the procedure, and was 19.0% [13, 25] when it was not; $prop_{diff}$ = .19 [.10, .26]. The imagery procedures used in implantation studies tend to be fairly elaborate. Participants are typically given instructions for context reinstatement and then guidance in imagining specific details of events. Thus we cannot determine exactly which aspects of these procedures contribute to false memory formation.

Providing photographs depicting suggested events. There is evidence that the medium in which the suggested event is presented can also influence the rate of false recall. For instance, Garry and Wade (2005) found higher rates of false memory formation when participants were presented with a description of the false event rather than a doctored photograph of the false event. Garry and Wade concluded that such photos may serve to constrain imagination during efforts to recall the event, thus leading to lower memory formation. We coded whether participants viewed (4 studies; N = 170) or did not view (6 studies; N = 253) a photo that allegedly depicted the specific suggested event. Consistent with their argument, weak but statistically reliable negative relationships emerged for Coherence (r = -.12), Imagery Amount (r

= -.19) and Imagery Quality (r = -.16). The false memory rate was 25.9% [95% CI 20, 32] when a photo depicted the event and 33.6% [28, 40] when no such photo was presented; $prop_{diff}$ = .08 [.00, .16]. One caveat worth noting is that only one event (hot air balloon ride) and the same or a similar photo has been used in all of the studies that have provided a doctored photo of the event.

Combining self-relevant information, imagination, and photos. To examine the combined power of these techniques (providing self-relevant information, encouraging imagining, presenting a suggested narrative without a photo depicting the event), we contrasted study conditions in which all three procedures were present (3 studies; N = 89) against all other observations (7 studies; N = 334). The false memory rate was 46.1% [95% CI 36, 56] when the three conditions were present and 26.3% [22, 31] for when they were not; $prop_{diff}$ = .20, [.09, .31]. As Figure 3 shows, the main difference is that participants in the studies that met those conditions were more often judged to have formed 'Full' false memories.

These three suggestive techniques — self-relevant information, imagination, and narratives without photos — might work in concert as follows: The presence of self-relevant information may enhance a participant's belief that the suggested event occurred, encouraging efforts to recall it. Self-relevant information may also provide participants with an effective memory cue. This may create an environment that is ripe for thoughts, images, and feelings to come to mind, and may bias participants toward accepting them as memories. Participants may embellish those mental experiences as they repeatedly imagine the suggested event. Moreover, if participants are provided with a verbal description of the suggested event, rather than a doctored photo that depicts precisely how the event occurred, they also have free rein to imagine. Together these processes may encourage participants to develop perceptually rich and vivid mental experiences that they ultimately misinterpret as memories of the suggested event. We note, however, that our

results must be interpreted with caution, because participants were not randomly assigned to these different suggestive conditions and there were confounding differences between the studies that did versus did not combine suggestive techniques.

False Autobiographical Belief

To this point, our examination of the data has emphasized false memories. However, many theories of false memory formation state that false autobiographical beliefs are a step along the path to the development of a false recollection. Coming to believe an event is a genuine part of the past is sufficient to lead to changes in views of the self and behavior (e.g., Mazzoni & Kirsch, 2002). The findings in Bernstein et al. (2015) indicate that once an event is sufficiently believed to have occurred, development of recollection does not contribute further to behavioral outcomes. In other words, belief in the occurrence of an event may be sufficient to influence behavior, whether or not there is also an accompanying episodic recollection.

Given the relevance of the development of false autobiographical belief to understanding the effects of suggesting events, we report here on subjective autobiographical belief ratings that were taken in two of the samples (N = 116; Desjardins & Scoboria, 2007; Hessen Kayfitz & Scoboria, 2012). The average belief in occurrence rating (1 to 8 point scale) following the implantation procedure across all individuals was 4.71 (95% CI 4.23, 5.19). The average belief rating did not differ statistically across the false memory categories discussed above (none, accepted, partial, full, robust, rejected). As can be seen in the frequency distribution in Figure 4, the distribution of scores was not normal. While 50% of individuals rated autobiographical belief above the scale midpoint, 28% reported autobiographical belief at the top of the scale. This pattern indicates that implantation procedures promote a notable degree of false belief formation in a majority of participants, and that a high level of belief is not uncommon following these

procedures. The level of autobiographical belief results from the procedure regardless of the degree of development of false recollection, which varies across the false-memory categories as discussed above in the analysis of reliving and remember-know ratings (see Scoboria, Wysman, & Otgaar, 2012, for further discussion).

Summary

In this mega-analysis, we systematically applied the same coding system to over 400 memory reports from previously published false memory implantation studies. This method afforded us the statistical power necessary to ask more refined questions that have only been speculated about in prior work. Setting methodological differences between the eight target studies aside, across studies 30.4% of participants were classified as having developed a false memory, and 53.3% were judged to have accepted the suggested event as genuine to some degree (inclusive of those classified as having developed a false memory). The presence of idiosyncratic self-related information, an imagination procedure during the suggestion, and to a lesser extent presenting the suggestion without a photo depicting the specific event, were each associated with high memory formation rates. Given all three of these conditions we found a false memory rate of 46.1% and an acceptance rate of 69.7%.

Consistent with models of false memory formation (Hyman & Kleinknecht, 1999; Mazzoni & Kirsch, 2002; Scoboria, Mazzoni, Kirsch, & Relyea, 2004), our analysis shows that if a false event is suggested, if evidence is provided that the event occurred, if resistance to considering the possibility that the event occurred can be overcome, and if imagination is employed, then false autobiographical memories often arise. Furthermore, suggestive practices appear to instill a degree of belief in false events in about two-thirds of study participants, and autobiographical belief ratings are often high following the procedure. Although implantation

studies have largely been concerned with illusory autobiographical memories, on an applied level *false autobiographical beliefs* may be more important. Bernstein et al. (2015) recently presented evidence that it may be changes in autobiographical belief that underpin changes in suggestion-related attitudes and behavioral intentions.

Our purpose here is not to review the extant body of research relevant to false memories, and we remind readers that the memory implantation methodology emphasized here is just one method in a rich body of research on memory errors. Distinct experimental methods, correlational studies, case reports, and anecdotal reports converge to show that people can come to remember false memories for personally meaningful past events. While the implantation method remains the most compelling experimental method for demonstrating and manipulating the formation of rich autobiographical false memories, a complete understanding of the processes that contribute to memory formation must be based in a wider reading of the literature. For that purpose, we direct readers to Schacter, Chamberlain, Gaesser, and Gerlach (2012), Laney and Loftus (2013), and Loftus (2005) for reviews.

One important finding here is that a memory report can look like a genuine memory to observers, even if the person does not explicitly report remembering the event. Indeed, 10% of participants in our combined sample appeared to have strong episodic mental representations of the suggested event, but ultimately reported that they did not remember it. This finding is important for at least two reasons. First, people may not experience recollection even though their reports meet coding criteria, and conversely, people may experience recollection when their reports do not meet the criteria (see also Otgaar, Scoboria, & Smeets, 2013). The disconnect between a participant's reported subjective experience and the criteria being applied can lead to the over- or under-reporting of false memory rates. Second, so-called objective judgments of

events as recollected may not coincide with participants' subjective experience of remembering. Given these complications, the elicitation and classification of reports (and the underlying cognitive states that contribute to them) must be examined carefully. Most memory implantation studies have mechanisms built in to ensure that false memory estimates are appropriate. For instance, many studies have distinguished between participants who report images only and those who report memories. In most studies, independent judges are trained to look for clear evidence of remembering, and disputes between judges are usually classified into the more conservative category. Finally, some studies, including this one, have documented convergence between the judges' memory categorizations and participants' self-ratings. Together, these procedures help to minimize the likelihood of coding error. Methodologically, our findings reinforce the importance of obtaining clear and elaborate statements from participants about what they remember and what they believe about suggested events (see Smeets, Telgen, Ost, Jelicic, & Merckelbach, 2009, for an excellent example of the value of clarifying memory claims).

One conclusion that can be made based on this study and on similar work is that it can be difficult to objectively determine when someone is recollecting the past, versus reporting other forms of knowledge or belief or describing mental representations that have originated in other sources of experience. Even under highly controlled laboratory conditions, memory researchers struggle to define and observe memory. How, then, can we expect therapists, forensic investigators, medical personnel, human resource staff, or jurists to be any better at this task?

Potential limitations

Study selection. We included published false memory implantation studies with adults conducted in English that we could obtain at the time that we conducted the project. We chose to include published studies that had passed the peer review process. We did not include studies

that examined the development of false memories in children due to the need to consider developmental differences with child participants, and because few studies that used the implantation procedure with children were available when we began the project. Beyond the eight studies included here, we were unable to obtain information from four published studies for which the researchers indicated that the transcripts no longer existed (Heaps & Nash, 2001; Hyman, Husband, & Billings, 1995; Loftus & Pickrell, 1995; Porter, Yuille, & Lehman, 1999). The total *N* from those studies was 221 (range 20 to 77); and the false memory rates reported ranged from 20% to 56% (weighted average 39.5%).

Of course, publication bias is likely to have affected the effect sizes we report here. We were not aware of any unpublished false memory implantation studies that met our criteria and for which transcripts existed at the time of our analysis. Nonetheless, to minimise selection bias, we included all of the participants who took part in the eight target studies in our analyses, regardless of the condition to which they were assigned. The target studies also varied in their intention to promote false memory development; for instance, some conditions served as controls and other conditions were expected to produce minimal development or even inhibit false memory formation. For this reason, one should not interpret the percentages of memories classified as partial, full, and robust as representing some limit in the level of false memory creation.

The retrieval of genuine events. Some participants in implantation studies who appear to have false memories may be retrieving genuine memories of past experiences (e.g., perhaps the participant really did experience a hot-air balloon ride). But there are several reasons to believe that such cases are rare. First, the critical events used in memory implantation studies are often the kinds of events that would have required parental permission or guidance, yet parents

indicate that the event never occurred. Other critical events are so peculiar and distinctive – such as putting slime in a teacher's desk (Lindsay et al., 2004) or spilling a bowl of punch on a bride (Hyman & Pentland, 1996) – that it is not plausible that many participants had experienced those events. Even if some participants had experienced a similar event the suggested false event is unlikely to perfectly match in terms of the exact age and manner with any genuine childhood experience. Second, several studies using different procedures have addressed this issue directly by showing that people can develop distorted beliefs and memories about highly implausible or impossible events (Braun, Ellis, & Loftus, 2002; Mazzoni, Loftus & Kirsch, 2001; Mazzoni & Memon, 2003; Spanos, Burgess, Burgess, Samuels, & Blois, 1999). One procedure used to address such concerns in memory implantation studies is to ask participants' family members whether, to their knowledge, the participants ever experienced the false event(s). All eight of the studies included in our analysis included this step for verifying that the false event never occurred. While it is an important step, we acknowledge that parents/siblings can only provide information that is, to the best of their knowledge, true. We cannot assume that just because an informant indicates that their child never experienced an event that in fact their child never had a related experience. Finally, research suggests that people incorporate details from a variety of cognitive processes, such as imagination, guided-imagery, and genuine memories, to construct memories of suggested events (Wade et al., 2002). It is simply incorrect to state that false memories are entirely false (Lindsay & Read, 1995; Read & Lindsay, 1994) and many false memories are likely to be built from elements of genuine experiences (Hyman et al., 1995).

We reiterate that a complete understanding of errors in autobiographical memory requires a broader reading of the literature. The memory implantation procedure focused on in this paper is complemented by laboratory-based methods in which researchers have control over the events

that are recalled. For example, Goff and Roediger (1996) described a method for inducing false memories for actions performed in the laboratory, and their approach has been further adapted by others (e.g., Clark, Nash, Fincham, & Mazzoni, 2013; Nash, Wade, & Lindsay, 2009; Seamon, Philbin, & Harrison, 2006; Thomas & Loftus, 2002). One important trade-off between the implantation procedure and laboratory procedures is ecological validity, because the events used in controlled laboratory studies do not resemble autobiographical memories in their full complexity. We also emphasize that our focus is on the development of false memories, and not on the issue of retrieval of true memories.

Implications

This research demonstrates the value of identifying situations in which formal or informal meta-analytic reviews are inappropriate for making summary statements about the state of knowledge within a literature. We offer an example of how to conduct integrative analyses in which measurement is reconceptualised in a theoretically informed way that allows data to be combined across diverse studies. In contrast, Brewin and Andrews (2016) recently reviewed false memory implantation research in what we perceive to be an example of what should not be done when attempting to summarize this literature. While they stated that they were not attempting to conduct a meta-analysis, they did report summary estimates of, for example, false memory formation rates. By averaging across studies and conditions, Brewin and Andrews concluded that the false memory formation rates across extant studies is no higher than 15%. They did acknowledge a number of potential moderating variables, but they then ignored the influence of such moderators on memory formation and, crucially, implied that their 15% estimate is a universal limit that generalizes to real-world cases. Our perspective is more modest and situationally grounded. Our estimate of the rate of false memories across the eight studies

analyzed here was substantially higher than 15%. But our broader point is that the likelihood of false memories depends on interactions between many variables. Under some conditions, false memories of various kinds are very rare, and under others they are very common. Psychologists are a long way from having a complete understanding of how the relevant variables interact to determine the likelihood and nature of false-memory experiences, but mega-analyses such as this may help us develop such an understanding. (For further discussion of Brewin and Andrew's analyses and conclusions, see Lindsay & Hyman, in press; Nash, Wade, Garry, Loftus, & Ost, in press; McNally, in press; Otgaar, Merckelbach, Jelicic, & Smeets, in press; Pezdek & Blandon-Gitlin, in press; Scoboria & Mazzoni, in press; Smeets, Merckelbach, Jelicic, & Otgaar, in press).

To conclude, research on false memory formation originated in criticisms of the notion that suggestive therapeutic practices were specifically and reliably linked to the recovery of genuine memories (Lindsay & Read, 2001). Despite a substantial body of evidence that suggestive practices promote the development of false memories, these notions remain prevalent in North America (Patihis et al., 2014) and in the UK (Ost et al., 2013). Our results firmly support the assertion that suggesting false events can produce false memory in a substantial percentage of people. On the one hand, it is likely to be more difficult to lead people to develop false beliefs or memories of childhood sexual abuse than of, say, a childhood prank. On the other hand, participants in the studies reviewed here were undergraduate students and the suggestive pressures brought to bear were brief and relatively mild. With stronger techniques combining the factors we investigated, more people may be led to create false memories of negative experiences. Moreover, people struggling with psychopathology who seek help for their symptoms may be particularly vulnerable to suggestions, and some trauma-memory-oriented 'treatments' use much more extensive procedures (psychopharmacology, hypnosis, pressure to

recall, elaborate explanations of memory repression and memory recovery, exposure to autobiographies in which memory has been ‘successfully recovered’, and so forth), often with greater frequency and over longer periods of time. We are hopeful that the use of dangerously suggestive practices has declined over the last two decades, but some treatment providers continue to use these practices (see Cara, 2014, for a recent case). Memory implantation research has already contributed to changing practice and policy by psychotherapists and by individuals who conduct forensic interviews. Our results reinforce how important is it to continue educating people about the malleability of memory.

References

Addis, D.R., Pan, L., Vu, M.A., Laiser, N., Schacter, D.L. (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia, 47*, 2222–2238. doi:10.1016/j.neuropsychologia.2008.10.026

Bernstein, D., & Loftus, E.F. (2009). How to tell if a particular memory is true or false. *Perspectives on Psychological Science, 4,* 370-374.

Bernstein, D., Scoboria, A., & Arnold, R. (2015). The consequences of suggesting false childhood food events. *Acta Psychologica, 156,* 1-7.

Bluck, S., Alea, N., Habermas, T., & Rubin, D. C. (2005). A tale of three functions: The self-reported uses of autobiographical memory. *Social Cognition, 23,* 91-117. doi: 10.1521/soco.23.1.91.59198

Braun, K.A., Ellis, R., & Loftus, E.F. (2002). Make my memory: How advertising can change our memories of the past. *Psychology and Marketing, 19,* 1-23.

Brewer, W. F. (1986). What is autobiographical memory? In D. Rubin (Ed.), *Autobiographical memory* (pp. 25-49). Cambridge: Cambridge University Press.

Brewer, W. F. (1996). What is collective memory? In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 96-66). New York: Cambridge University Press.

Brewin, C.R., & Andrews, B. (in press). Creating memories for false autobiographical events: A systematic review. *Applied Cognitive Psychology.* doi: 10.1002/acp.3220

Brown, N. R., Schweickart, O., & Svob, C. (2016). The effect of collective transitions on the organization and contents of autobiographical memory: A transition-theory perspective. *American Journal of Psychology, 129*, 259-282.

Cara, E. (2014). The most dangerous idea in mental health. Pacific Standard (magazine, November 13, 2014). Retrieved on August 17, 2015 from: http://www.psmag.com/health-and-behavior/dangerous-idea-mental-health-93325.

Christianson, S., & Safer, M.A. (1996). Emotional events and emotions in autobiographical memories. In D. C. Rubin (Ed), (1996). *Remembering our past: Studies in autobiographical memory*, (pp. 218-243). New York, NY: Cambridge University Press

Clark, A., Nash, R.A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. PLoS ONE, 7, e32998. doi: 10.1371/journal.pone.0032998

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261.

Conway, M.A., & Loveday, C. (2015). Remembering, imagining, false memories & personal meanings. *Consciousness & Cognition*, *33*, 574-581.

Cumming, G. (2013). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York: Routledge.

de Maat, S., Dekker, J., Schoevers, R., van Aalst, G., Gijsbers-van Wijk, C., Hendriksen, M., ... & de Jonghe, F. (2008). Short psychodynamic supportive psychotherapy, antidepressants, and their combination in the treatment of major depression: a mega-analysis based on three randomized clinical trials. *Depression and Anxiety*, *25*, 565-574.

Desjardins, T., & Scoboria, A. (2007). "You and your best friend Suzy put slime in Ms. Smollett's

desk": Producing false memories with self-relevant details. *Psychonomic Bulletin & Review, 14*, 1090-1095. doi:10.3758/BF03193096

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin and Review, 8,* 155-161.

Fivush, R., Habermas, T., Waters, T. E., & Zaman, W. (2011). The making of autobiographical memory: intersections of culture, narratives and identity. *International Journal of Psychology, 46,* 321-45. doi: 10.1080/00207594.2011.596541.

Garry, M., & Wade, K. A. (2005). Actually, a picture is worth less than 45 words: narratives produce more false memories than photographs do. *Psychonomic Bulletin and Review, 8,* 155-161.

Ghetti, S., & Alexander, K. W. (2004). “If it happened, I would remember it”: Strategic use of event memorability in the rejection of false autobiographical events. *Child Development, 75,* 542-561. doi: 10.1111/j.1467-8624.2004.00692.x

Goff, L. M., & Roediger, H. L., III (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition, 26,* 20-33.

Greenberg, D.L., & Knowlton, B.J. (2014). The role of visual imagery in autobiographical memory. *Memory and Cognition, 42,* 922-934. doi: 10.3758/s13421-014-0402-5.

Greenwald, A. G., & Banaji, M. R. (1989). The self as a memory system: Powerful, but ordinary. *Journal of Personality and Social Psychology, 57,* 41-54.

Hallahan, B., Newell, J., Soares, J., et al. (2011). Structural MRI in Bipolar disorder: an international collaborative mega-analysis of individual adult patient data. *International Clinical Psychopharmacology, 26,* e64.

Heaps and Nash (2001). Comparing recollective experience in true and false autobiographical memories. *Journal of Experimental Psychology, 27,* 920-30.

Hessen Kayfitz, J., & Scoboria, A. (2012). False memory is in the details: Photographic details predict memory formation. *Applied Cognitive Psychology, 26,* 333-341. doi: 10.1002/acp.1839

Hyman, I. & Billings, F.J. (1998). Individual differences and the creation of false childhood memories. *Memory, 6,* 1-20.

Hyman, I. E., Husband, F., & Billings, J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology, 9,* 181-197. doi: 10.1002/acp.2350090302

Hyman, I. E., Jr., & Kleinknecht, E. (1999). False childhood memories: Research, theory, and applications. In L. M. Williams & V. L. Banyard (Eds.) *Trauma and memory* (pp. 175-188). Thousand Oaks, CA: Sage.

Hyman, I. E., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language, 35,* 101-117. doi: 10.1006/jmla.1996.0006

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391-422). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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,* 371-376. doi:10.1037/0096-3445.117.4.371

Johnson, M. K., Hastroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28. doi: 10.1037.0033-2909.144.1.3

Johnson, M.K., & Sherman, S. (1990). Constructing and reconstructing the past and the future in the present. In T. E. Higgins & R. M. Sorrentino (Eds), *Handbook of motivation and cognition: Foundations of social behavior,* Vol. 2. , (pp. 482-526). New York, NY, US: Guilford Press, xviii, 621 pp.

Kassin, S. (2012). Why confessions trump innocence. *American Psychologist, 67,* 431-445. doi: 10.1037/a0028212

Lampinen, J. M., & Odegaard, T. N. (2006). Memory editing mechanisms. *Memory, 14,* 649-654. doi: 10.1080/09658210600648407

Laney, C., & Loftus, E.F. (2013). Recent advances in false memory research. *South African Journal of Psychology, 43,* 137-146.

Lindsay, D. S. (2008). Source monitoring. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory*. *Vol. 2 of Learning and memory: A comprehensive reference*, 4 vols. (J. Byrne, Editor) (pp. 325-347). Oxford: Elsevier.

Lindsay, D. S., Hagen, L., Read, J. D., Wade, K., & Garry, M. (2004). True photographs and false memories. *Psychological Science, 15,* 149-154. doi: 10.1111/j.0956-7976.2004.01503002.x

Lindsay, D.S., & Hyman, I.E. (in press). Commentary on Brewin and Andrews. *Applied Cognitive Psychology*.

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

Lindsay, D. S., & Read, J. D. (2001). The recovered memories controversy: Where do we go from here? In G. M. Davies & T. Dalgleish (Eds.), *Recovered memories: Seeking the middle ground* (pp. 71–94). London: Wiley.

Loftus, E. F. (1993) The reality of repressed memories. *American Psychologist*, *48*, 518-537.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*, 361-366.

Loftus, E. F., & Bernstein, D. M. (2005). Rich false memories: The royal road to success. In A.F. Healey (Ed.), *Experimental Cognitive Psychology and its Applications* (pp 101-113). Washington DC: American Psychological Association.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25,* 720-725.

Mazzoni, G., & Kirsch I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 121-145). Cambridge, UK: Cambridge University Press.

Mazzoni, G. A. L., Loftus, E. F., & Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: A little plausibility goes a long way. *Journal of Experimental Psychology: Applied, 7,* 51-59. doi: 10.1037/1076-898X.7.1.51

Mazzoni, G., & Memon, A. (2003). Imagination can create false autobiographical memories. *Psychological Science, 14,* 186-188. doi: 10.1046/j.1432.1327.1999.00020.x

McNally, R. (in press). False memories in the laboratory and in life: Commentary on Brewin and Andrews (2016). *Applied Cognitive Psychology.*

Nash, R, Wade, K., Garry, M., Loftus, E.F., & Ost, J. (in press). Misrepresentations and flawed logic about the prevalence of false memories. *Applied Cognitive Psychology.*

Nash, R. A., Wade, K. A., & Lindsay, D. S. (2009). Digitally manipulating memory: Effects of doctored videos and imagination in distorting beliefs and memories. *Memory and Cognition, 37,* 414-424. doi: 10.3758/MC.37.4.414

Neisser, U. (1987). A sense of where you are: Functions of the spatial module. In P. Ellen, and C. Thinus-Blanc (Eds.), *Cognitive processes and spatial orientation in animal and man, Vol. 1: Experimental animal psychology and ethology* (pp. 293-310). Dordrecht, Netherlands: Martinus Nijhoff Publishing.

Neisser, U. (1994). Self-narratives: True and false. In U. Neisser & R. Fivush (Eds.), *The Remembering Self: Construction and Accuracy in the Self-Narrative*. Cambridge University Press: Cambridge.

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

Olkin, I. (1995). Statistical and theoretical considerations in meta-analysis. *Journal of Clinical Epidemiology, 48,* 133-146.

Ost, J., Easton, S., Hope, L., French, C. C., & Wright, D. B (*in press*). Latent variables underlying beliefs about memory among Chartered Clinical Psychologists, Hypnotherapists and undergraduate students. *Memory*.

Ost, J., Foster, S., Costall, A., & Bull, R. (2005). False reports of childhood events in appropriate interviews. *Memory, 13,* 700-710. doi: 10.1080/09658210444000340

Ost, J., Wright, D., Easton, S., Hope, L., & French, C. (2013). Recovered memories, satanic abuse, Dissociative Identity Disorder and false memories in the UK: a survey of clinical psychologists and hypnotherapists. *Psychology, Crime & Law, 19,* 1-19

Otgaar, H., Merckelbach, H., Jelicic, M., & Smeets, T. (in press). The potential for false memories is bigger than Brewin and Andrews suggest. *Applied Cognitive Psychology*.

Otgaar, H., Scoboria, A., & Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 717-730. doi: 10.1037/a0029668

Patihis, L., Ho, L. Y., Tingen, I. W., Lilienfeld, S. O. & Loftus, E. F. (2014). Are the "Memory Wars" over? A scientist-practioner gap in beliefs about repressed memory. *Psychological Science, 25,* 519-530.

Pezdek, K., Blandon-Gitlin, I., Lam, S., Hart, R. E., Schooler, J. W., (2006). Is knowing believing? The role of event plausibility and background knowledge in planting false beliefs about the personal past. *Memory & Cognition, 34,* 1628-1635.

Pezdek, K., & Blandon-Gitlin, I. (in press). It is just harder to construct memories for false autobiographical events. *Applied Cognitive Psychology.*

Pezdek, K., Finger, K., & Hodge, D. (1997). Planting false childhood memories: The role of event plausibility. *Psychological Science, 8,* 437-441.

Pillemer, D. B. (1998). *Momentous events, vivid memories.* Cambridge, MA, US: Harvard University Press.

Pillemer, D. B. (2003). Directive functions of autobiographical memory: The guiding power of the specific episode. *Memory, 11,* 193-202.

Porter, S., Yuille, J. C., & Lehman, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: implications for the recovered memory debate. *Law and Human Behavior, 23,* 517-537.

Read, J. D., & Lindsay, D. S. (1994). Moving toward a middle ground on the 'false memory debate': reply to commentaries on Lindsay and Read. *Applied Cognitive Psychology, 8,* 407–435.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96*, 341-357. doi:10.1037/0033-295X.96.2.341

Rubin, D. C. (2005). The basic-systems model of episodic memory. *Perspectives on Psychological Science, 1,* 277-311. doi: 10.1111/j.1745-6916.2006.00017.x

Pezdek, K., & Blandon-Gitlin, I. (2016). It is just harder to construct memories for false autobiographical events. *Applied Cognitive Psychology.*

Schacter, D. L. (2013). Memory: sins and virtues. *Annals of the New York Academy of Sciences, 1303*, 56-60.

Schacter D. L., Chamberlain, J., Gaesser, B., & Gerlach, K. D. (2012). Neuroimaging of true, false, and imaginary memories: findings and implications. In: Nadel L., Sinnott-Armstrong W. (Eds.), *Memory and law: Perspectives from cognitive neuroscience*. New York: Oxford University Press; In Press.

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

Scoboria, A., & Mazzoni G., (in press). Invited commentary on Brewin and Andrews (2016). *Applied Cognitive Psychology*.

Scoboria, A., Mazzoni, G., Jarry, J., & Shapero, D. (2012). Implausibility inhibits but does not eliminate false autobiographical beliefs. *Canadian Journal of Experimental Psychology, 66,* 259-267. doi: 10.1037/a0030017

Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology, 18,* 791-807. doi: 10.1005/acp.1062

Scoboria, A., Talarico, J., & Pascal, L. (2015). Metamemory appraisals in autobiographical event recall. *Cognition, 136,* 337-349. 10.1016/j.cognition.2014.11.028

Scoboria, A., Wysman, L., & Otgaar, H. (2012). Credible suggestions affect false autobiographical beliefs. *Memory, 5,* 429-442.

Seamon, J. G., Philbin, M. M., & Harrison, L. G. (2006). Do you remembering proposing marriage to the Pepsi machine? False recollections from a campus walk. *Psychonomic Bulletin & Review, 13,* 752-756.

Smeets, T., Merckelbach, H., Jelicic, M., & Otgaar, H. (in press). Dangerously neglecting court realities. *Applied Cognitive Psychology.*

Smeets, T., Telgen. S., Ost, J., Jelicic, M., & Merckelbach, H. (2009). What's behind crashing memories? Plausibility, belief and memory in reports of having seen non-existent images. *Applied Cognitive Psychology, 23,* 1333-1341. doi: 10.1002/acp.1544

Spanos, N.P., Burgess, C.A., Burgess, M.F., Samuels, C., & Blois, W.O. (1999). Creating false memories of infancy with hypnotic and non-hypnotic procedures. *Applied Cognitive Psychology, 13,* 201-218.

Sternberg, K.J., Baradaran, L.P., Abbott, C.B., Lamb, M.E., & Guterman, E. (2006). Type of violence, age, and gender differences in the effects of family violence on children's behavior problems: A mega-analysis. *Developmental Review, 26,* 89-112.

Strange, D., Wade, K., & Hayne, H. (2008). Creating false memories for events that occurred before versus after the offset of childhood amnesia. *Memory, 16,* 475-484. doi: 10.1080/09658210802059049

Suddendorf, T. & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic Social and General Psychology Monographs, 123,* 133–167.

Szpunar, K. K., Addis, D. R., McLelland, V. C., Schacter, D. L. (2013). Memories of the future: New insights into the adaptive value of episodic memory. *Frontiers in Behavioral Neuroscience*, *7*, 47.

Talarico, J.M., LaBar, K.S., & Rubin, D.C. (2004). Emotional intensity predicts autobiographical memory experience. *Memory and Cognition, 32,* 1118-1132.

Thomas, A. K., & Loftus, E. F. (2002). Creating bizarre false memories through imagination. *Memory & Cognition, 30,* 423-431.

Tulving, E. (1983). *Elements of Episodic Memory.* Oxford: Clarendon Press.

Wade, K. A., Garry, M. (2005). Strategies for verifying false autobiographical memories. *The American Journal of Psychology, 118,* 587-602.

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,* 597-603. doi: 10.3758/BF03196318

Wuensch, K. L. (2012). *Using SPSS to obtain a confidence interval for Cohen's d*. http://core.ecu.edu/psyc/wuenschk/SPSS/CI-d-SPSS.pdf.

Table 1. Definitions of categories in false memory implantation studies conducted with adults.

Study	Category	Definition (Quotations taken from published papers)
Loftus & Pickrell (1995)	Full false memory	[no definition provided]
	Partial false memory	"partial memories included remembering parts of the event and speculations about how and when it might have happened"
*Hyman & Pentland (1996); [Hyman, Husband, & Billings (1995); Hyman & Billings (1998)]	Clear false memory	"reports of spilling the punch, consistent elaborations, and statements that the event was a memory"
	Partial false memory	"consistent elaborations with some statements of remembering but did not include memory of actually spilling the punch"
	Trying to remember	"described an image or reported related self-knowledge, but made no clear claims to remember the event"
Pezdek, Finger & Hodge (1997)	Remembered	"An event was operationally defined as remembered if the subject recalled specific details of the event that were not included in the description read"
Porter, Yuille, & Lehman (1999)	Implanted memory	"reported remembering the suggested event; …agreed with and/or incorporated the information clues into the memory report; …reported more information than the four information clues (prompts)"
	Partially implanted	"recalled information or imagery pertaining to the event, but never recalled it in its entirety or was unsure whether the memory was real"
Heaps & Nash (2001)	False memory	"given a rating of 1-4 on the following question by participants… How much of the event do you remember? (0 None, 1 little, 2 some, 3 most, 4 all), and subsequently remembered content could not be verified by parents as having happened"
*Wade, Garry, Read, & Lindsay (2003)	Clear false memory	"report memories of the critical balloon ride, including consistent elaboration of information not depicted in the photograph"
	Partial false memory	"consistently elaborated… but did not indicate memories of taking the balloon ride per se"
	Trying to remember	"tried to recall the false event and described images of it but did not claim these images as memories"
*Lindsay, Hagen, Read, Wade, & Garry (2004); [*Desjardins & Scoboria (2007); *Hessen-Kayfitz & Scoboria (2012)]	False memories	"if the subject appeared to believe that s/he was remembering the suggested event"
	Images but not memories	"described images associated with the suggested event but did not appear to experience those images as memories of the event"
*Garry & Wade, 2005	False memory	"report remembering the false event and details about the event that were not depicted in the doctored photograph or outlined in the false narrative"
	Images only	"described images associated with the false event, but those images need not have been experienced as memories"
*Ost, Foster, Costall, & Bull (2005)	Full memory	"reported remembering the suggested event; …agreed with and/or incorporated the information clues (prompts) into the memory report; …reported more information than the information clues (prompts)"
	Partial memory	"reported that they remembered the event but, for example, provided no more information than had been provided in the prompts, or indicated that they were unsure it was a real memory…"
*Strange, Wade, Hayne (2008)	False memory	"…claimed to remember the event and reported at least two specific details about it"
	Image	"…speculated about at least three different aspects of the event"
Otgaar, Scoboria, & Smeets (2013)	False memory	"Participants who indicated 'remembering' the event and provided additional event-related details"

Note: In all of the studies, cases that did not meet one of the definitions were included in a 'no memory' category. * Indicates studies included in the current analysis.

Table 2. Items used to code memory transcripts

Item	Scale	Definition
Acceptance of suggested details	Which of the following best describes the participant's acceptance of the specific details of the event that were provided to them, at the end of the entire process? 0 Outright rejection; 1 Minimal acceptance / Accept parts, reject parts; 2 Moderate acceptance / Acceptance with no active rejection; 3 Complete acceptance	Consider the specific details that were suggested to the participant. To what degree do they accept just these details at the end of the entire process? Do they accept all of the material as accurate, do they accept some parts but not other parts, do they struggle with accepting it, are they unsure?
Elaboration	Does the participant provide any more information beyond what was provided to them? No / Yes [If Yes] Elaboration amount: How much elaboration is provided? 1 Minimal; 2 Moderate; 3 Substantial Elaboration related: How closely related is the additional information to the information provided? 0 Not related; 1 Minimally related; 2 Moderately related; 3 Highly related	Does the participant include any additional information beyond the original information provided to them when describing the event (the material suggested)? (examples: describing what occurred before or after, adding new characters, new objects, new scenes, new actions, adding details to the information provided, etc.)
Imagery	Does the participant generate and/or experience any sensory images of the event? No / Yes [If Yes] Imagery amount: How many sensory images does the participant experience? 1 Small amount of imagery (about 1 image); 2 Moderate amount of imagery (about 2-4 images; 3 High amount of imagery (about 5 or more images) Imagery quality: What is the quality of the best sensory image experienced?1 Weak imagery; 2 Moderate imagery; 3 Strong (vivid) imagery	When participants envision the event, do they experience any images? That is, do they appear to be perceiving the way things looked, sounded, smelled, or felt?
Coherence	How coherent is the participant's description of the event (how well does it hang together as an organized narrative)? 0 Not coherent; No information; 1 Minimally coherent; 2 Somewhat coherent; 3 Highly coherent	How organized and coherent is the participant's description of the event (no matter the amount of information that is described)? Does it make logical sense? Does it flow? Or is their description broken up into chunks that don't necessarily fit together? Are there gaps between the reported series of events? Rate the entire transcript as a whole.
Emotional content	Does the participant include any emotional content when discussing the event? No / Yes [If Yes] Emotion amount: How much emotional content is present? 1 Little emotional content (about 1 reference); 2 Moderate emotional content (about 2-4 references); 3 High emotional content (about 5 or more references) Emotion engagement: Which of the following describes the emotional engagement in whatever amount of emotional content is present?0 No engagement; 1 Weak engagement; 2 Moderate engagement; 3 Strong engagement	Definition: Does the participant include emotional material in their description of the past event? (eg., "I felt…", "We were very excited…", "I might have felt…", "Someone in that situation would probably feel…" etc.). If there is any use of emotional state (worry, sad, happy, loved, hurt, enthusiastic, rage, anger, etc.), emotional expression (yelled, screamed, laughed, cried, zany), or other related words such as "feeling", "mood", "temper", and so forth, then this should be coded as "Yes".
Judged memory	Do you think that the participant feels they have a memory for the specific event? 0 No; 1 Unsure; 2 Yes, part; 3 Yes, for entire event	Considering all of the information available to you, do you feel that the participant has a memory for the event? Partial means that some but not all of the provided information is endorsed as remembered; Full means that all of the provided information is endorsed as remembered.
Event rejection	Does the participant state that they do not have a memory / do not remember (or similar) at the end of the procedure? No / Yes	

Table 3. Correlations between independent judges' ratings for key memory features and the judged memory variable.

	Acceptance	Elaboration	Imagery	Judged memory, partial or full	Judged memory, full
Acceptance present	1.00	.56 [.49,.63]	.50 [.41,.57]	.54 [.47,.60]	.39 [.32,.46]
Elaboration present		1.00	.73 [.65,.81]	.35 [.29,.40]	.18 [.12,.23]
Imagery present			1.00	.38 [.33,.43]	.21 [.16,.26]

Note: Each feature was coded as absent or present by a different pair of raters. Spearman's rho [95% confidence intervals]. N = 423. 'Judged memory' was coded as partial memory or full memory by one pair of raters.

Table 4. Definitions of ‘memory’ and associated mean ratings on other judged variables.

	Acceptance	Elaboration	Imagery	N	Emotion intensity Mean [95% CI]	Coherence Mean [95% CI]	Judged memory Mean [95% CI]
1	None	-	-	153	0.06 [0.02,0.11]	0.16 [0.09, 0.24]	0.03 [0.00, 0.08]
2	Any	-	-	270	0.45 [0.36,0.55]	1.33 [1.22, 1.46]	1.09 [0.97, 1.23]
3	Moderate	-	-	173	0.60 [0.49,0.71]	1.65 [1.51, 1.78]	1.43 [1.29, 1.58]
4	Moderate	Any	Any	137	0.63 [0.51,0.76]	1.72 [1.58, 1.86]	1.49 [1.34, 1.65]
5	Moderate	Mod	Mod	91	0.86 [0.70,1.03]	2.08 [1.93, 2.23]	1.86 [1.65, 2.06]
6	High	Mod	Mod	44	1.07 [0.83,1.32]	2.30 [2.12, 2.48]	2.05 [1.76, 2.35]

Note: Each line indicates a different combination of the Acceptance, Elaboration and Imagery judged variables. Starting with line 2, all subsequent lines are a sub-set of the preceding group (e.g., the 173 on line 3 are a subset of the 270 on line 2). Range for all variables is 0 to 3. See Table 5 for the Ns associated with each of the final groups. Cases unique to Line 4 were labeled “Partial false memory”; those unique to Line 5, “Full false memory”; and those on Line 6, “Robust false memory”.

Table 5. Descriptive statistics for remaining judged items by coded memory category.

	Judged variable					
Category	Emotion amount	Emotion intensity	Coherence	Elaboration relationship	Imagery quality	Judged memory
No memory (N = 153)	0.12 [0.07,0.19]	0.07 [0.03,0.11]	0.17 [0.10,0.25]	0.65 [0.48,0.81]	0.42 [0.32,0.52]	0.03 [0.00,0.08]
Accept (N = 97)	0.33 [0.20,0.46]	0.14 [0.07,0.25]	0.63 [0.47,0.79]	1.48 [1.24,1.72]	0.97 [0.81,1.14]	0.36 [0.25,0.51]
Partial false memory (N = 36)	0.64 [0.36,0.91]	0.28 [0.11,0.47]	1.19 [.94,1.44]	2.70 [2.51,2.86]	1.31 [1.14,1.53]	1.22 [.86,1.56]
Full false memory (N = 46)	0.98 [0.72,1.22]	0.65 [0.41,0.89]	1.91 [1.65,2.19]	2.64 [2.45,2.81]	1.98 [1.76,2.19]	1.86 [1.67,2.06]
Robust false memory (N = 47)	1.30 [1.02,1.60]	1.09 [0.80,1.39]	2.40 [2.23,2.57]	2.87 [2.60,2.91]	2.00 [1.79,2.19]	2.38 [2.10,2.62]
Rejected (N = 44)	0.75 [0.47,1.03]	0.36 [0.18,0.56]	1.25 [1.02,1.49]	2.54 [2.34,2.75]	1.50 [1.32,1.70]	0.00 [0.00,0.00]

Notes: All scales ranged from 0 to 3. Means and 95% confidence intervals.

Table 6. Correlations between judged categories and suggestive techniques.

Judged item	Idiosyncratic self-relevant information provided in suggestion	Suggested event imagined during retrieval procedure	Doctored photo depicts suggested event
Emotion amount	.09 [-0.01,0.18]	**.24** [0.15,0.31]	.02 [-0.07,0.11]
Emotion engagement	.02 [-0.07,0.12]	**.21** [0.12,0.28]	-.05 [-0.14,0.04]
Coherence	**.23** [0.14,0.33]	**.38** [0.30,0.46]	**-.12** [-0.21,-0.02]
Imagery present	**.26** [0.17,0.35]	**.24** [0.15,0.33]	-.10 [-0.19,0.00]
Imagery amount	**.30** [0.21,0.39]	**.26** [0.16,0.35]	**-.19** [-0.28,-0.10]
Imagery quality	**.28** [0.19,0.36]	**.28** [0.20,0.36]	**-.16** [-0.24,-0.07]
Elaboration present	**.25** [0.16,0.34]	**.23** [0.14,0.33]	-.09 [-0.17,0.02]
Elaboration amount	**.19** [0.11,0.29]	**.27** [0.18,0.36]	**-.14** [-0.23,-0.05]
Elaboration relationship	**.16** [0.06,0.25]	**.20** [0.10,0.29]	-.06 [-0.15,0.04]
Accept event	**.19** [0.10,0.29]	**.27** [0.18,0.36]	.01 [-0.10,0.10]
Judged memory	**.16** [0.07,0.26]	**.21** [0.12,0.30]	**-.11** [-0.20,-0.01]

Notes: Study characteristics coded as 0 Not present, 1 Present. Coefficients are Spearman's rho. N = 423; error bars are 95% confidence intervals based on bootstrapping 1000 samples. Correlations which do not include zero in the confidence interval are in bold. Range for all variables is 0 to 3.

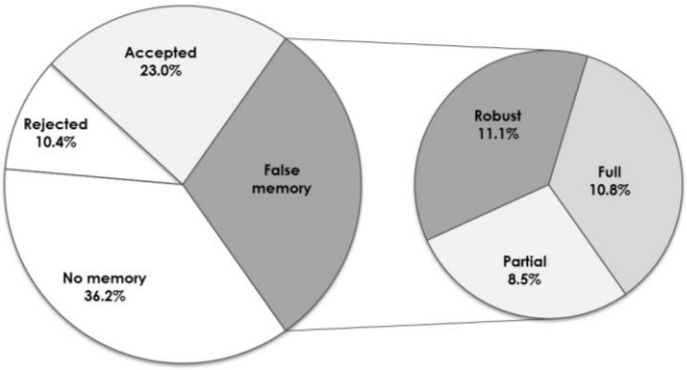

Figure 1. Final categorization of 423 participants' memory reports.

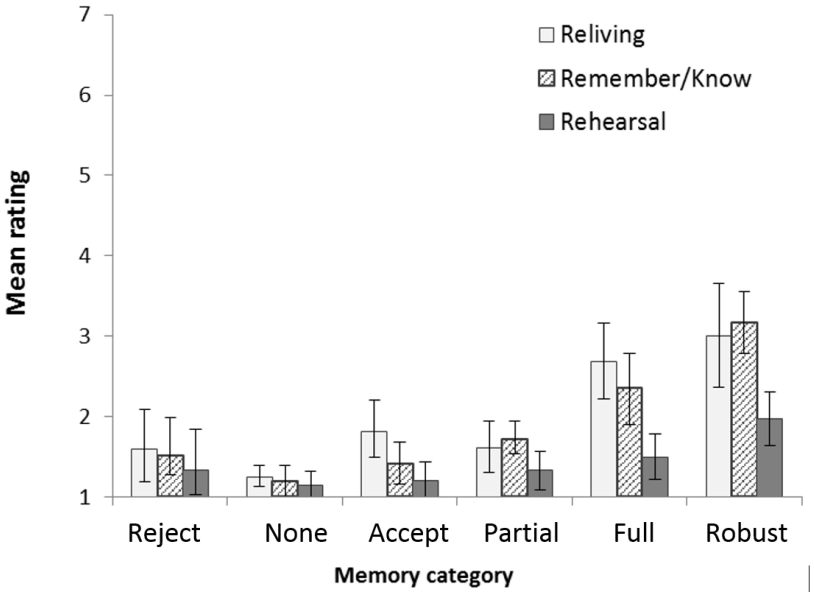

Figure 2. Participants' mean subjective ratings by false memory category. Remember/know and rehearsal ratings were available for 75.9% of the combined sample; Reliving ratings were available for 58.6% of the sample. Ns by group for the remember/know and rehearsal items are: 30, 123, 67, 34, 39, 28. Ns by group for the reliving item are: 17, 106, 43, 24, 39, 19. Error bars show 95% confidence intervals on group means.

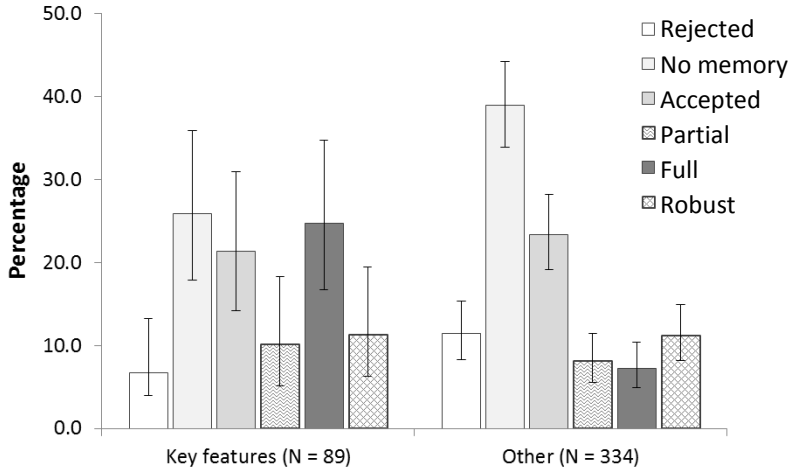

Figure 3. Percentage of memory reports within each memory category for cases with (left bars) and without (right bars) key suggestive techniques (idiosyncratic information, imagination, a suggestive narrative that was not accompanied by a photo depicting the suggested event). Error bars show 95% confidence intervals on proportions.

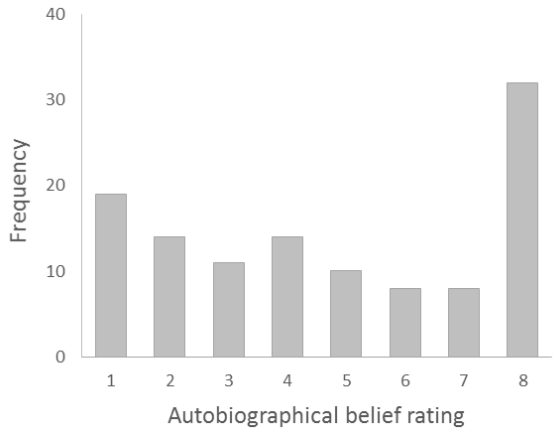

Figure 4. Frequency of autobiographical belief ratings for suggested events at the end of the studies on the 8-point scale (*n* 116).