

Crashing Memory 2.0: False Memories in Adults for an Upsetting Childhood Event

Lawrence Patihis^{1*} and Elizabeth F. Loftus²

¹University of Southern Mississippi; ²University of California, Irvine

Corresponding author: Lawrence Patihis, Department of Psychology, Owings-McQuagge Hall,
University of Southern Mississippi, Hattiesburg, MS 39406. Email: L.Patihis@usm.edu

ACKNOWLEDGEMENTS

We thank Steven Frenda for help at the design and data collection stages the study, and research assistants Marlene Lin Ma, Anita Chen, Joseph McCall, Matthew Miller, Lavina Ho, Stephany Debski, Stephanie M. Martinez, Zahra Bozorgkhan, Kenneth Nguyen, Patricia Place, Gina Machiaverna, Alexandra Chindris, Samuel Cretcher, Monica Aguilar, Paddy Asgari, Yesenia Orozco, Luis Garcia, Natalie Ross, Devan Duenas, Howard Lee, and Hillary Patton for help during data collection.

Abstract

Previous crashing memory studies have shown that adults can be led to believe they witnessed video footage of news events for which no video footage actually exists. The current study is the first to investigate adults' tendency to report memories of viewing footage that took place when they were children: the plane crash in Pennsylvania on September 11, 2001. We found that in a computer questionnaire 33% indicated a false memory with at least one false detail. In a more detailed face-to-face interview, only 13% of the group described a detailed false memory. Familiarity with the news story, fantasy proneness, alcohol use, and frequency of negative emotions after 9/11 were all associated with a persistent false memory. Participants who had received prior suggestion were more likely to later report false memories in the subsequent interview. We discuss our novel results and the importance of the paradigm.

Keywords: false memory, crashing memory, childhood, memory distortion false belief,

Crashing Memory 2.0: False Memories in Adults for an Upsetting Childhood Event

Memory distortion research has important applications in real-life cases in the law, clinical psychology, and other areas. Some of these real-life cases have involved the distortion of memory in children, in adults for recent events, and in adults for events in their childhood. To investigate both applied and theoretical issues related to such real life phenomena, researchers have developed a number of memory distortion paradigms, including the misinformation effect (Loftus, Miller, & Burns, 1978), associative word lists (Roediger & McDermott, 1995), rich false memory (Loftus & Pickrell, 1995), and imagination inflation (Garry, Manning, Loftus, & Sherman, 1996). One paradigm, the "crashing memory" technique (see Crombag, Wagenaar, & van Koppen, 1996), involves asking participants if they have seen footage for a widely reported news event, often involving a crash (hence the name), when in fact no video footage of the event really exists. In response to such questions, many participants in these studies appeared to develop false beliefs and memories that they witnessed events they could not possibly have seen.

The crashing memory paradigm has several advantages that represented an important alternative way to measure memory distortion. The first advantage is that the event involved a nationally important news story that was usually personally important to the participants, thus allowing for the measurement of an upsetting memory that can be autobiographical in nature. The second advantage is the national news event in question is somewhat of a collective experience across participants, and thus their memory reports are comparable across subjects. This is often not true in other studies where people are asked about their upsetting or traumatic memories and different participants point to completely different types of events in their lives. The third advantage is that the researcher can be reasonably confident that the event (viewing

disturbing footage in this case) did not take place. This certainty about the falsity of the memory is not always present in other autobiographical false memory research (e.g., Loftus & Pickrell, 1995, Garry et al., 1996). Despite these advantages, previous crashing memory studies have focused on events that happened when the participants were adults. In the present study we explore the phenomenon by asking adults about an event that occurred in their childhood. We also investigate a number of other factors that have not previously been examined in crashing memory studies, such as the effect of prior suggestion.

Previous Crashing Memory Studies

In the first crashing memory study, Crombag et al. (1996) told Dutch participants there was videotape of a widely reported Boeing 747 crash into apartments in Amsterdam. Although the actual plane crash had not been filmed, a remarkable 55% in Study 1 and 66% in Study 2 of the participants reported seeing the footage and about 45% reported details of its contents. A follow up study demonstrated a similar phenomenon for a nationally important car crash. Ost, Vrij, Costall, and Bull (2002) asked participants if they had seen footage of the car crash in which Princess Diana was killed (when in fact none exists), and 44% of the sample reported they had.

The studies that followed reinforced these surprisingly large percentages with different target news events, and also showed that social influence could boost false memory rates. For example, Granhag, Stromwall, and Billings (2003) found that 55% of respondents reported that they saw nonexistent footage of a well-known incident involving a sinking ferry, and found that about 30% gave a false detail, suggesting that the false beliefs may have been accompanied by memories. These results showed that misleading comments from peers, when overheard by the participant, can boost false memory rates (see also Ost, Hogbin, & Granhag,

2006 for a replication using a different target news event). Similarly, Wilson and French (2006) asked participants to recall the details of a bombing that had occurred in a Bali nightclub, and 36% of participants reported seeing nonexistent footage, with nearly all of them reporting corresponding memories for details they could not have seen.

These studies raised the question whether these false memory reports were being caused by the suggestion, or whether they were spontaneously generated. An interesting experimental approach by Smeets et al. (2006) found that the level of suggestiveness in the crashing memory questions affected the rates at which participants reported seeing nonexistent news footage of the assassination of Dutch politician Pim Fortuyn. Interestingly, even with no suggestion (“Do you remember whether there was a film...”; notice the indefinite article “a”), 27% of participants indicated they had seen the footage of the assassination, although only 6% reported actual false details. With false suggestion (“Did you see the amateur film of the Fortuyn shooting?”; notice the definitive article “the”) the false reports were much higher: 63% with 33% giving false details. These results demonstrate the importance of suggestive wording, but also that false memories can occur even in the absence of misleading post-event information (spontaneously; cf. Mazzoni, 2002).

Further research established the importance of familiarity of the news event to the participant. Ost, Granhag, Udell, and Hjelmsäter (2008) asked participants, 150 from Sweden and 150 from the United Kingdom to complete questionnaires about the explosion of the No. 30 bus in Tavistock Square, London. United Kingdom participants were more likely to say they had seen nonexistent computer-generated image of the explosion, and nonexistent television footage of the explosion, compared to the Swedish participants (40% vs. 16%). This

indicated that perhaps familiarity with the event, or indeed the amount of exposure to the original news reports, might facilitate the implanting of such memories.

A study by Smeets et al. (2009) perhaps created doubts about how many of the false memory reports in previous crashing memory studies were in fact valid false memories that would persist. In keeping with a number of past studies, they found high rates of false report of seeing nonexistent footage (66%) following the suggestion that there is footage of the assassination of Pim Fortuyn (a political figure well known to the Dutch participants). However, after the debriefing, in which they were informed that there is no footage of the actual moment of the assassination, most claimed they had not fully understood the question, with only 10% maintaining that they had truly remembered the footage during the experiment. This decrease in false memory percentages in a crashing memory study raised questions whether previous studies could have had inflated percentages, and also what in particular could have caused the reduction in percentages. In Smeets et al., it is unclear whether the *true* post-event information (i.e. ‘there is no footage’) caused the decrease in false memory reports or whether the open ended and more detailed, clear, and cautious approach of the post-debriefing interview caused the decrease. Ost et al. (2008) and Smeets et al. (2009) raised the issue of whether the suggestive question in crashing memory studies simply brings out a pre-existing false memory (spontaneous; formed before the study, perhaps by exposure to successive media reports) or the suggestion in the experiments causes the false memory production (suggestion-dependent; cf. Mazzoni, 2002).

Individual differences and crashing false memory. Some previous research has been done investigating individual differences and crashing memories. In some studies, females were more prone to false report (Crombag et al., 1996; Jellic, Smeets, Candel, van Suijdam, &

Merckelbach, 2006a, Jelicic et al., 2006b) but not in other studies (Ost et al., 2002, Ost et al., 2006; Granhag et al., 2003, Smeets et al., 2009). With regards to personality related measures, Ost et al. (2008) found that those scoring high on dissociation (Dissociative Experiences Scale, DES-C; Wright & Loftus, 1999) or fantasy proneness (Creative Experiences Scale, CEQ; Merckelbach, Horselenberg, & Muris, 2001) were more likely to give definitive details of nonexistent footage (of the bus moving in the London bus bombings). These measures relate to an ongoing debate as to whether the most prone people who develop false memories are those which are highly hypnotizable (related to fantasy proneness), a question that has historical roots in false memory production in hypnosis (see Patihis & Younes Burton, in press). The question of whether dissociation is related to false memories is routed on an ongoing debate about the relationship, if any, between dissociation, and memory (e.g. Lynn et al., 2014). Another area of interest is the possibility that participants' tendency towards social desirability might explain the unusually high frequency of false memory reports in previous crashing memory research (cf. McCloskey & Zaragoza, 1985). And finally, an individual difference that may also affect memory is alcohol consumption, which has shown promise as a factor in other areas of memory distortion research (e.g., Garfinkel, Dienes, & Duka, 2006) but there is a scarcity of data addressing alcohol in crashing memory studies. In the current study we investigate the role of these variables, and others.

The Current Study

Although previous research seemed to establish a strong effect of suggestion on memory of nonexistent footage, Smeets et al. (2009) raised the question as to how many of these are genuine and indeed persistent false memories. They also discussed an apparent reduction in false memories, and the reasons for it. A key question raised by this was whether false memories in

crashing memory studies were spontaneously produced by prior exposure to media (and merely exposed, rather than caused by, by the suggestion in the experiments) or whether they were caused by the suggestion in the actual experiments. The current study explores these issues. Previous crashing memory studies have also not examined events that took place in the childhood of the participants. The current study does this to gain an insight to the real-life parallel of false memories induced into adults for upsetting childhood events. We also investigate a number of possible correlates of these autobiographical false memories: such as gender, fantasy proneness, and dissociation.

Method

Participants

The participants were 297 undergraduates (221 females; 78 males) who participated in the study in exchange for course credit. Of these, 48.8% self-identified as Asian/Indian, 21.5% as Caucasian, 16.2% as Hispanic/Latino, 7.4% as Middle Eastern, 3.4% as Hawaiian/Pacific Islander, 2.0% as African American / Black, and .7% other. Participants' mean age was 20.1 years ($SD = 3.1$). The majority of our participants ($n = 271$) were between 18 and 22 so that they were between the ages of 8 and 12 at the time of the event. For purposes of comparison there were also a small number ($n = 26$) of older participants, between ages 23 to 51.

Design

Participants were randomly assigned to one of two conditions¹: an experimental condition (in which they completed a computer questionnaire containing suggestive information about the target event), or a control condition (in which they completed a computer questionnaire about an unrelated event). Approximately 40 minutes after the questionnaires, all participants underwent a

face-to-face interview in which they were asked about their memories for the target event. See Figure 1 for a graphical representation of the study design.

[INSERT FIGURE 1 ABOUT HERE]

Materials and Procedure

Participants were invited to participate in a study ostensibly about “Personality, Individuality, and Slideshows.” All subjects participated in 2011, between March and November, approximately 10 years after 9/11. Participants came into the laboratory one at a time, staggered one every 15-20 minutes, and were greeted by a research assistant who gave them verbal instructions to prepare them for the study. Between one and three participants participated in a computer laboratory room at any given time, with one or two research assistants supervising. The lab room was windowless and mostly silent. The research assistants who interacted with the participants and conducted the interview were blind to both the assigned condition of participants and the precise hypotheses of the study.

Session 1. Once randomly assigned to condition, the participant proceeded to fill out the computer questionnaires. They first answered demographic questions. Other measures asked about memory for their negative emotions in the week following September 11, 2001. In addition to those pretest questions, participants also completed a number of measures, including an alcohol-use scale (modified from LaBrie, Hummer, Grant, & Lac, 2010), the Creative Experiences Scale (fantasy proneness; Merckelbach, Muris, & Rassin, 1999; Merckelbach, Horselenberg, & Muris, 2001), and the Dissociative Experiences Scale (DES-C; Wright & Loftus, 1999), . Session 1 typically took participants about 35 minutes to complete.

Session 2. Exactly one week after Session 1, the subject returned to the lab for Session 2. As in Session 1, a research assistant gave verbal instructions and then helped the participant get started on the computer questionnaire.

Crashing memory questionnaire. Depending on which condition they were assigned to, participants either completed a news story questionnaire asking about their memory for United 93 crash footage (experimental condition), or the Human Genome news event (control group). Each questionnaire was similar in construction and length and differed only in the memory recall target (i.e. the targets were United 93 crash *footage*, versus the Human Genome *news-event*). See the Supplemental Materials for the crashing memory materials used in the computer questionnaire.

United 93 condition. In this condition, participants were told that footage of the crash exists and has been widely shown, and were then asked whether they had seen the footage. They were then asked to indicate details of the footage, which involved force choice questions involving false details, including an option of “I don’t remember”. For example, one question asked about the participants memory of the video footage, and gave the choices that the plane (a) “came down vertically, nose down and almost without forward speed”, (b) “slid into the ground almost horizontally and at considerable speed”, or (c) “I can’t remember.” Another question gave the participants the choice that the video footage was (a) “very clear, you can see and hear exactly what is happening,” (b) “fuzzy, it is difficult to tell what is happening,” or (c) “I can’t remember.” For the full wording of the questionnaires, see Supplemental Material. Participants then proceeded to fill out a Memory Characteristics Questionnaire (MCQ, modified from Johnson, Foley, Suengas, & Raye, 1988 to include an “ I don’t remember” option, scored as 0)

about their memory for the footage. This condition acted as the experimental condition, simulating the crashing memory manipulations of previous studies.

Control condition. In this condition neither United 93 nor September 11, 2001 were mentioned at all. Instead, participants were asked a similar set of questions about their memory for the Human Genome *news-event* (as opposed to footage) that was reported in 2001. This condition acted as a control condition that both had no suggestion of nonexistent footage, but also did not stimulate any imagery of United 93 or 9/11. These two conditions allowed us to measure the effect of prior misleading suggestion on the interview which occurred approximately 40 minutes later.

Fillers. Subjects completed a number of filler tasks that for reasons of clarity and focus are not analyzed in the current study. Participants also completed a 91-item personality scale designed to measure a total of 13 personality subscales (Swedish Universities Scale of Personality, SSP; Gustavsson, et al., 2000). These trait subscales include somatic trait anxiety, psychic trait anxiety, stress susceptibility, lack of assertiveness, impulsiveness, adventure seeking, detachment, social desirability, embitterment, trait irritability, mistrust, verbal trait aggression, and physical trait aggression. The social desirability scale, in particular, was used to investigate the participants' wish to help the researchers (i.e. demand characteristics) as a possible confound. In addition, research assistants who conducted the interviews also completed this personality scale as well as demographic questions (for reasons of space and clarity, the results for interviewer characteristics are given in the Supplemental Materials).

United 93 crash footage interviews. In this audio-recorded structured interview, participants were taken away from the other participants into a sound proof room. In a one-on-one conversation with the research assistant, face-to-face, participants were told about the

various crashes on 9/11, then told that we are focusing here only on United 93, and asked “Are you familiar with this event?” (open response). The participants are then told that there is footage for the crash of United 93, and then asked “Do you remember seeing that footage?” (open response). No such footage actually exists. In this interview, we took pains to be clear about what crash we were referring to, and reiterated it was a crash into a *field*. See Supplemental Materials for the interview script. Those assigned to the United 93 and the Human Genome control condition participated in the interview. In the interview, those participants who said "yes" they had seen the footage were then asked follow up questions about details. After the interview was complete we revealed that the study was actually about memory distortions, and that there is no footage of the United 93 crash. The typical duration of Session 2 was between 60 and 75 minutes.

Quantitative Coding of Interview Responses

All interviews were coded by two independent coders (research assistants), and any inter-rater disagreements were scrutinized carefully and resolved by a supervising researcher in discussion with one or both research assistants. The question asking about whether the participant was familiar with United 93 was coded as 0 = no, .5 = unsure/maybe, and 1 = yes, and the initial inter-rater agreement rate was 78% (66 disagreements out of 297; Cronbach α = .839). Whether someone was familiar with the United 93 news event was not always easy to code, because sometimes participants would start out by saying "a bit" but then go on to give details to demonstrate that they were fully familiar. These difficult-to-code cases were carefully recoded whenever research assistants' initial coding did not match. The questions asking whether they had seen the footage were coded in a similar manner (no = 0; maybe/unsure = .5; yes = 1) and the initial inter-rater agreement rates on those questions were 93% and 92% respectively

(Cronbach $\alpha = .955$ and $.935$). The question asking how well the participants remember the video on a scale from 1 to 10 was straightforward to code, and the inter-rater agreement rate was 99% (Cronbach $\alpha = .996$). Interviews that included a “yes” response *and* a memory score of more than 2 on the 1–10 Likert scale were coded for false details of the actual crash by a supervising researcher (LP). All of these 43 cases examined had the coding justified by transcribing the relevant part of the interview text (done by a research assistant or the supervising researcher) and considering whether the actual content of the detail *must* be false (i.e. movement of the plane or the actual impact/explosion was coded as a false detail, whereas pictures or video of the aftermath were not assumed to be false). Details from any of the other crashes on 9/11 (e.g. a crash into a building) were not coded as a false detail.

Results

In the results that follow, we examine the rates of false memory in the questionnaire, and then the interview. We then examine the association between false memory and familiarity with the news-story, prior suggestion, ethnicity, alcohol use, emotion, and fantasy proneness. Finally, we take a specific look at those participants who were between the ages of 8 and 12 at the time of the target event (9/11).

Computer Questionnaire

Of the 297 participants completing the study, 95 were assigned to the Human Genome control no-suggestion questionnaire², and 202 were assigned to the Crash Footage questionnaire (for an explanation of the unequal group sizes, see Footnote 1). Of the 202 participants assigned to the United 93 Crash Footage condition, when asked "have you seen the video?" 36.6% indicated "yes," they had seen the United 93 crash footage. Of the 36.6% who indicated "yes," they had seen the footage, 91.8% gave at least one false detail (how the plane moved, clarity of

footage, or length of footage). Specifically, of those 36.6% who said "yes," 62.2% gave a detail about the plane moving, 60.8% gave a detail about after the impact, 68.9% gave some detail about the clarity of the footage, and 51.4% gave a detail about the length of the footage (see Supplemental Material for question wording on the 4 detail questions). Of *all* the 202 participants, 41.1% gave some false detail of the nonexistent plane crash footage (therefore a minority of 4.5% participants initially indicated they had not seen the footage, but consequently chose a false detail in the follow-up questions).

Of the 202 participants in the group that received false suggestion (the Crash Footage condition), when asked "how well can you remember having seen the video?" 59.4% indicated 1 (*no memory at all*) on the scale from 1 to 10, and 40.3% indicated a 2 or above. 9.9% indicated a score of 5 or above, with one participant (.5%) indicating 10 (*a very clear memory*). Of those 73 that had indicated "yes" they had seen the footage, all (73; 100%) indicated a score of 2 or above on this scale, with 26% (19) indicating a score of 5 or above. In contrast, of the 128 who indicated "no" they had not seen the video, 93.8% (120) indicated a score of 1 (*no memory at all*) on this 1–10 scale.

False Belief versus False Memory. We classified subjects as having a *Questionnaire False Memory* if all three of the following conditions were met: they indicated "yes" they had seen the footage, indicated a score of 2 or above on the aforementioned 1–10 scale, *and* chose at least one false detail. A *false belief* involved choosing "yes" they had seen the footage, but either choosing 1 (*no memory at all*) on the 1–10 scale or indicating no false details in the follow up questions. Using this criteria, 67 (33.2%) of the 202 participants indicated a Questionnaire False Memory, 7 (3.5%) indicated a false belief, and 128 (63.4%) reported no false belief or memory.

Audio Recorded Interview

About 40 minutes after the computer questionnaire, those in the United 93 Crash footage (Experimental, $n = 202$) condition, and the Human Genome (Control, $n = 95$) condition participated in the face to face recorded interview with a research assistant. Of the 297 participants who participated in the crashing memory interview, 48 (16.2%) said they had seen the footage, 23 (7.7%) said unsure/maybe, and 226 (76.1%) said they had not seen crash footage of United 93. This figure of 16.2% saying “yes” in the interview compares to 36.6% that indicated “yes” in the forced-choice computer questionnaire earlier. In the interview, when asked how well they remembered having seen the video on a scale from 1 to 10, 37.0% (110 out of 297) indicated 1 (*no memory at all*) on the scale from 1 to 10, 63.0% indicated a 2 or above, and 15.8% indicated a score of 5 or above, with one participant (.3%) indicating 10 (*a very clear memory*).

False Memory versus False Belief. Similarly to the questionnaire, we categorized participants as having an *Interview False Memory* when all the following three conditions were met in the interview: the participants says “yes” they have seen the video, they give a score of 2 or above on the aforementioned 1–10 scale, *and* they give some false detail from the footage connected to the actual crash (most commonly being how the plane moved or crashed in the footage). A *false belief* involves the participants saying “yes” they had seen the footage, but then either choosing 1 (*no memory at all*) on the 1–10 scale *or* giving no false details of the actual crash in the footage. Using these criteria we identified 30 (10.1%) as having an Interview False Memory, 18 (6.1%) having a false belief, and 249 (83.8%) showing neither.

Comparing Experimental to Control Group

Effect of Prior Suggestion on Interview False Memory. Some subjects who participated in the audio recorded interview had responded to a suggestive questionnaire about

United 93 on the computer 40 minutes earlier (Experimental condition), while others had been randomly assigned to a non-suggestive neutral questionnaire about the 2001 human genome news story (Control condition). Figure 2 shows that a higher proportion of those in the Experimental condition exhibited Interview False Memories (13.4%; 27 of 202) compared to those in the Control condition (3.2%, 3 of 95), $\chi^2(1, N = 297) = 7.42$, exact $p = .006$, Cramer's $V = .158$. This four-fold difference in False Memory proportions in the interview shows the potent effect of the suggestive questionnaire 40 minutes earlier.

[INSERT FIGURE 2 ABOUT HERE]

Effect of Prior Suggestion on Quality of Interview False Memories. Within those who demonstrated an Interview False Memory, we examined whether there is a difference in memory quality between those who had prior suggestion 40 minutes earlier ($n = 27$), and those who hadn't ($n = 3$). We found that those with prior suggestion scored marginally higher on the question "how well can you remember having seen the video on a scale from 1 to 10, where 1 means no memory at all and 10 means a very clear memory" (No prior suggestion: $M = 2.33$, $SD = 0.58$; Prior Suggestion: $M = 4.09$, $SD = 1.48$), $t(28) = 2.01$, $p = .054$, Cohen's $d = 1.57$. Levene's test for equality of variances had a p -value of .15 ($F = 2.24$) indicating that assumption of equality of variance is not *strictly* violated. However, the low Levene's test p -value and the small group size warrant reporting the equal variances *not* assumed test: comparing prior suggestion to no prior suggestion on subsequent false memory clarity in the interview yielded $t(5.76) = 4.01$, $p = .008$.

Memory Characteristics Questionnaire (MCQ) and Interview False Memories. Forty minutes before the interview, those 202 participants in the experimental condition answered MCQ items relating to the United 93 footage. Most of the items on the MCQ associated with the

occurrence of an Interview False Memory (r s from .13 to .41; see Supplemental Material Table S1 for each MCQ item and effect sizes). The largest MCQ predictor of Interview False Memory was item 10 “Feelings at the time that I first saw the video were” (Anchors: 1 = *not intense*, 7 = *very intense*), $r = .41$, $p < .001$. In other words, self-reported memory of more intense feelings was associated with a greater likelihood of reporting a detailed false memory in the interview.

Persistent False Memory. We defined a *Persistent False Memory* as being when a participant met the three criteria for a Questionnaire False Memory *and* the three criteria for an Interview False Memory. By this metric, 25 (12.4%) participants exhibited a Persistent False Memory, out of the 202 who participated in both the suggestive questionnaire and interview. To clarify apparent differences in percentages, this Persistent False Memory figure of 12% is by definition from within the experimental group only, whereas the aforementioned 10% Interview False Memory is from the whole sample (some of whom did not get the prior suggestive questionnaire).

Comparing those Familiar to those Not Familiar with the News Story

Although all participants were familiar with the 9/11 general news story, only 142 (70.3%) of the 202 in the Experimental condition were somewhat familiar with the United 93 story (responses to question “are you familiar with this event?” coded .5 “maybe,” $n = 20$; or 1 “yes,” $n = 122$). In the *questionnaire*, the 142 participants rated as being familiar with the United 93 news story, compared to those 60 not familiar (coded 0 = “no”), were marginally more likely to indicate a Questionnaire False Memory (37.3% vs. 23.3%; $\chi^2(1, N = 202) = 6.50$, $p = .054$, exact $p = .072$ (all p s two-tailed throughout), Cramer’s $V = .136$). In the *interview* those somewhat familiar with the United 93 news story, compared to those not familiar, were more likely to indicate an Interview False Memory (19% vs. 0.0%; $\chi^2(1, N = 202) = 13.2$, exact $p <$

.001, Cramer's $V = .255$). Similarly, those somewhat familiar with the United 93 news story, compared to those not familiar, were more likely to indicate a Persistent False Memory throughout the questionnaire and interview (17.6% vs. 0.0%; $\chi^2(1, N = 202) = 12.06$, exact $p < .001$, Cramer's $V = .244$).

Media Exposure on 9/11. A bivariate logistic regression showed that self-reported news exposure on the *day of 9/11* was a marginal predictor of Persistent False Memory (Wald = 3.34, $p = .068$). Similarly, a bivariate logistic regression revealed that news exposure in the *month after 9/11* was a predictor of Persistent False Memory (Wald = 4.89, $p = .027$). However, putting both these media exposure variables into the same model resulted in neither being significant predictors of Persistent False Memory ($ps > .316$) which may be explained by the high inter-correlation between the two media variables ($r = .562$, $p < .001$) although the Variance Inflation Factor was not excessive (VIF = 1.67).

Individual Differences as Correlates of Persistent False Memory

We use Persistent False Memory as the main outcome variable in most of the analyses that follow because we consider it the cleanest and most conservative measure of a real visual false memory. If participants indicated a false memory in the questionnaire but not the interview it is possible they did not actually visualize a false memory, perhaps misunderstanding the question. There is also some doubt cast on Interview False Memories if the participant had previously indicated they had not seen the footage in the questionnaire. Only 2 participants in the experimental group gave a false memory in the interview after not indicating a false memory in the questionnaire³. Therefore, Persistent False Memory is only a slightly more conservative measure than Interview False Memory (12.4% vs. 13.4% in experimental group). In addition, using Persistent False Memory eliminates the possible confound of some participants having

prior suggestion and some not: all participants received prior suggestion. For these reasons, and in order to reduce the effects of multiple comparisons that would occur if we used several outcome measures, we chose Persistent False Memory as the best outcome measure for exploring most individual differences measures.

Demographics. There was no association between those who had a Persistent False Memory and gender, age, or on whether the participant was a psychology major or not ($ps > .332$). Specifically on gender, a variable discussed in previous research, females had statistically similar Persistent False Memory rates (12.5%) as did males (12.0%), $\chi^2(1, N = 202) = .009, p = .926$ (exact $p = 1.000$), Cramer's $V = .007$.

Those identifying their ethnicity as Middle Eastern had significantly higher rates of Persistent False Memory (46.7%) compared to those of other ethnicities (9.6%), $\chi^2(1, N = 202) = 17.57$, exact $p = .001$, Cramer's $V = .295$. This association between Middle Eastern ethnicity and Persistent False Memory was not effected by familiarity with the United 93, and was only slightly partially mediated by high-arousal negative emotion in the week following 9/11 (before mediation: $\beta = 2.11, p < .001$; after mediation: $\beta = 1.97, p = .001$)

Reported Emotion in Week After 9/11. Those who demonstrated a Persistent False Memory reported higher ratings of how often they had felt high arousal negative emotions in week after following 9/11 ($M = 4.12, SD = 2.78$), compared to those not exhibiting a Persistent False Memory ($M = 3.14, SD = 2.12$), $t(197) = 1.66, p = .040$, Cohen's $d = .40$. This report of their emotions in the week after occurred approximately 10 years after 9/11. This difference was less pronounced in negative emotions not categorized as high arousal ($ps > .133$).

Alcohol Use. Those who demonstrated a Persistent False Memory reported higher frequency of consuming alcohol than did those who did not show a Persistent False Memory (M

= 2.80: just below the anchor *twice a month*, $SD = 1.53$), compared to those not exhibiting a Persistent False Memory ($M = 2.14$, just above the anchor *once a month*, $SD = 1.39$), $t(200) = 2.20$, $p = .029$, Cohen's $d = .45$. This Likert scale was a fully anchored scale from 1 = *never* to 9 = *everyday* (from LaBrie, Hummer, Grant, & Luc, 2010). Similarly, those with a Persistent False Memory reported more heavy alcohol drinking binges (defined as 4/5 drinks in a two hour period) in the last two months ($M = 2.76$, $SD = 2.05$) compared to those who did not show a Persistent False Memory ($M = 1.94$, $SD = 1.59$), $t(200) = 2.33$, $p = .021$, Cohen's $d = .45$. This Likert scale ranged from 1 = *none*, to 9 = *10 or more times*. A summed composite of these and two other alcohol use questions also revealed a statistically significant difference, $t(200) = 2.01$, $p = .046$, Cohen's $d = .28$.

Fantasy Proneness and Dissociation. Those with a Persistent False Memory were significantly higher on fantasy proneness ($M = 11.48$, $SD = 4.25$; Creative Experiences Questionnaire) compared to those with no Persistent False Memory $M = 9.22$, $SD = 4.15$), $t(200) = 2.54$, $p = .012$, Cohen's $d = .54$. There were no differences between those with Persistent False Memory and those without on dissociation ($p = .953$; DES-C).

Personality. We found no association between Persistent False Memory and participants' personality trait subscales (SSP: see Method): social desirability, somatic trait anxiety, psychic trait anxiety, stress susceptibility, lack of assertiveness, impulsiveness, adventure seeking, detachment, embitterment, trait irritability, mistrust, verbal trait aggression, and physical trait aggression (all $ps > .201$). The Supplemental Materials document that we found no associations between the personality of the interviewer and false memory rates in the interview; with gender, age, and SES of the interviewer also yielding no effect (cf. Porter, Birt, Yuille, & Lehman, 2000).

Examining Those Ages 8–12 at the Time of the Target Event (9/11)

We replicated the above analysis excluding participants from age 23 to 51 ($n = 26$) because the majority of our participants were between 18 and 22 ($n = 271$) and we wanted here to focus on participants between the ages of 18 and 22 who were in mid-childhood (ages 8 to 12) when the target events of 9/11 occurred. Almost all the comparisons had similar outcomes as with the full sample. In other words, in those who were between 8 and 12 at the time of the target event (9/11), we found similar percentages of false memories as with the full sample. We also confirmed that relationships between familiarity, ethnicity, alcohol use, fantasy proneness, and false memories, existed in this 18–22 year old subgroup, as well as the full sample. However, in the 18–22 year old subgroup, we found all subtypes of memory for negative emotions in the week following 9/11 predicted false memories *generally*, whereas in the full sample high-arousal emotions were a stronger predictor than other types of emotion (see Supplemental Material for full analysis).

[INSERT FIGURE 3 ABOUT HERE]

Figure 3 shows the percentages of Persistent False Memory by age category at the time of the event. A Chi-Square analysis comparing Persistent False Memory Rates of those who were 12 or younger (13.1%; consolidating first five age categories in Figure 1) to those 13 or older at the time of the event (5.3%) revealed no significant difference, $\chi^2(1, N = 202) = .98, p = .323$, Cramer's $V = .070$. A Chi-Square analysis on all the 6 categorized age groups revealed no main effect for age group, $\chi^2(5, N = 202) = 4.83, p = .437$, Cramer's $V = .155$. This result was similarly not statistically significant even when excluding those age 23 or above, $\chi^2(4, N = 183) = 3.66, p = .454$, Cramer's $V = .141$. Similarly, age analyzed as a continuous variable was not

reliably associated with Persistent False Memory (full age range: $r_{bp} = -.069, p = .332$; limited 8–12 age range at time of 9/11: $r_{bp} = .017, p = .819$).

Qualitative Analysis of Interview False Memories

In the Supplemental Material we document some of the wording used by participants in the interview who demonstrated an Interview False Memory. In those examples it is clear that the participants understand which crash the interviewer is referring to. The United 93 crash is clearly distinguished from the other 9/11 crashes at the beginning of the interview (see the full interview script in the Supplemental Materials). Similar to the first interview excerpt, in the latter interview the emerging false memory seems relatively fragile at this stage, and although some false details are emerging (e.g. “burst into flames when it crashed into the floor”) it is also clear that these details are not yet filled out into a vivid and heavily detailed memory. This pattern was observed in many of the interview transcripts that demonstrated a false memory. It is because of this qualitative examination of the words used by participants in the interview that we conclude that the false memories are still in the fragile early stages of formation and not yet fully developed into multi-detailed and vivid autobiographical memory.

Discussion

The study produced false memory reports of seeing the nonexistent footage of the United 93 crash in a sizable minority of subjects in both the computer questionnaire, and in our more rigorous follow-up interview. This interview involved free recall (as opposed to forced-choice in the questionnaire) and we asked follow up questions to help us distinguish between a false memory and a false belief. We were conservative in that we required three prerequisite necessary conditions (a “yes”, a score not including 1 = *no memory at all* on a 1–10 scale, and a false detail of the actual crash) when distinguishing a false memory from a false belief in both the

questionnaire and the interview. Even with the necessary condition of a false detail accompanying the false report, we still found 12.4% had a Persistent False Memory throughout the questionnaire and the interview. Of those familiar with the United 93 news story, 17.6% had a Persistent False Memory, with media exposure marginally associating with such false memories. Prior suggestion 40 minutes earlier in the questionnaire had an additive effect on False Memory rates in the interview, with approximately a fourfold difference in percentages. Middle Eastern participants had higher rates of Persistent False Memory than people of other ethnicities. High arousal negative emotion, alcohol use, and fantasy proneness were associated with higher Persistent False Memory.

Despite many previous studies showing co-occurrence of a memory report and a false detail, some uncertainty about the paradigm was introduced by Smeets et al. (2009). That study raised the question of whether the crashing memory paradigm was actually producing false memories, as opposed to mistaken beliefs or other types of errors. In Smeets et al. (2009) they found that after participants were told there was actually no footage, 80% of those who had previously indicated a false memory retracted, many claiming that they had misunderstood the question. After debriefing, only 10% of the total sample in Smeets et al. (2009) indicated that they had had a false memory, compared to 66% before debriefing. In our current study, we also encountered a drop in the number of false memories when we explicitly clarified the events and the question in our face to face interview (from 33% to 13% within the experimental group). Our study shows that this reduction in false memory rates can be achieved before debriefing, which is to say that the reduction in our study was not caused by the post-event *true* information (as could have been a partial cause of the reduction in Smeets et al., 2009). The reduction in false memory rates in the current study is perhaps associated with the interview being clearer, more cautious,

and more detailed compared to the briefer questionnaire. One could also speculate that the face to face nature of the interview in our study elicited more caution and attention in the participants, compared to the anonymous clicking of the mouse on the computer questionnaire (cf. Koriat & Goldsmith, 1996). From this, we concur with Smeets et al. (2009) that even though there is a reduction of false memory rates, there is still a minority of people who exhibit what appears to be a genuine false memory. It is clear from both these studies that a good proportion of the emerging false memories are initially fragile and labile, because many initial false reports are corrected in the presence of true post-event information or more cautious questioning. However, it is possible that even some of these labile false reports would consolidate into full and stable false memories in the presence of reinforcement and time, and in the absence of debriefing or caution. Indeed, in the present study we show that two exposures to misleading information led to higher false memory rates than just one.

On a related note, Ost et al. (2008) and Smeets et al. (2009) raised the question whether false memories observed in crashing memory studies were already formed prior to suggestion (spontaneous) or caused by the suggestion in the experiment itself. Our use of random assignment during the questionnaire phase into groups that either received suggestion or no suggestion allowed us to conclude that suggestion in the experimental condition is an important cause of the false memories. This is evidence that perhaps a good proportion of false memory reports in crashing memory studies are in fact suggestion-dependent rather than spontaneous.

Qualitative evidence for this fragility of these newly formed false memories is perhaps gained from examining the actual wording used by the participants in the interview. It appears that even in those demonstrating false memories; those false memories are newly formed with only a few details. They appear relatively labile, and this fragility may explain why previous

research found that many false reports were retracted after correcting post-event information (Smeets et al., 2009), and why our percentages dropped from the questionnaire to the interview. It is for these reasons that we suspect that repeated reinforcement and time would be needed to solidify the false memories into elaborate and highly detailed accounts. This is supported by our finding that prior suggestion seemed to also lead to clearer false memories, compared to those with no prior suggestion. In addition, past research into implanting childhood false memories has shown that repetition (e.g. Ceci, Huffman, Smith, & Loftus, 1994) and repetition combined with suggestion (e.g. Ceci, Loftus, Leichtman, & Bruck, 1994) might be important for producing elaborate false memories that become richer in detail and more resistant to correction, compared to the early stages of false memory formation.

As in previous crashing memory studies (Ost et al., 2008) we found that fantasy proneness was associated with false memory. This can be explained by the ability of fantasy prone individuals to visualize and imagine events, and a tendency to remember a lot of visual details from the past (both accurately and inaccurately, see Merckelbach, 2004; cf. Patihis et al., 2013). This capacity likely helps bring the event to mind in a visual form, which in turn leads to the later consolidation of a full false memory. However, unlike Ost et al. (2008) we did not find that dissociation was associated with false memory production. Other personality traits also did not associate with false memories in the current study. In particular it is interesting that those who were high on social desirability had similar rates of false memory as those low on social desirability. This tends to dampen the idea that participants may be demonstrating false memories out of a wish to be socially accepted by the researchers.

Interestingly, we found that participants identifying as Middle Eastern had higher false memory rates, compared to other ethnicities. It is unclear why this might be. One possibility was

that because of the political connection of the terrorist attacks on 9/11 to the Middle East, those identifying as Middle Eastern might be more familiar with the news story. This, however, is probably not the case because those identifying as Middle Eastern had similar familiarity with the United 93 news event as other participants. Another possibility is that they had more negative emotions during and following 9/11: but this idea had only marginal support: memory of negative emotion only partially explained the relationship in a mediation analysis. For these reasons, though this result seems intuitively to make sense, it is difficult to say for certain the reason for the relationship between ethnicity and false memory for 9/11. In addition, caution is warranted here because only a small sample of 22 identified as Middle Eastern.

Interestingly, those who reported remembering more frequent experience of high-arousal negative emotions in the week following September 11th, 2001, had a higher proportion of false memories for the United 93 crash footage. This may be an indication that high emotional arousal may not only aid in memory consolidation (see Cahill & McGaugh, 1998) but paradoxically can lead to errors in details, in particular a higher susceptibility to false memories for some details (cf. Christianson & Loftus, 1987; also see the Paradoxical Negative Emotion hypothesis, Porter, Taylor, & ten Brinke, 2008).

There are some limitations to this study. Our measure for how often they felt negative emotions in the week after 9/11 would not be as accurate as if we had asked them soon after the event. This possible inaccuracy in this variable may come from the passage of time, or changing of how they appraise the events of 9/11 (cf. Levine, Whalen, Henker, & Jamner, 2005). One method for getting at this issue would be to conduct a longitudinal design, and future research could attempt that. In addition, this study is different than many previous studies in that a sizable

minority of participants was not familiar with the specific United 93 news story. Nevertheless, this allowed us to confirm that familiarity does indeed increase false memory rates.

This current study contributes to our current knowledge of false memory and crashing memory research in a number of ways. Firstly, it shows that reductions in false memory rates can be found by being more cautious and detailed in the way you ask the crashing memory questions, even after misleading suggestion has been presented. The current study establishes such a reduction in rates before debriefing, whereas Smeets et al. (2009) did so after also telling the participants that no footage actually exists. It also demonstrates that not all false memories in the crashing memory paradigm are so fragile that they either disappear with cautious further probing, or turn out to be a mere belief. We establish here that a minority are persistent and genuine detailed false memories. We argue that in the absence of debriefing that these semi-autobiographical false memories could persist, especially if reinforced. We show that repeated suggestion increased false memory rates in the crashing memory paradigm. We also identify factors that might increase these types of autobiographical memories, such as high-arousal emotions, alcohol use, prior suggestion, and familiarity. We also present some evidence that is congruent with previous findings showing a link between fantasy proneness and false memories. This study, and the crashing memory paradigm as a whole, gives memory researchers the opportunity to measure memory distortion in the autobiographical domain with the rare knowledge that we are sure the original event did not take place. News events such as 9/11 are important because they represent memories that can subjectively be experienced as flashbulb-like. Such collective experience represents the part of our autobiographical memory that is shared with millions of others. As Neisser (1982, p. 48) wrote, these are events in which we “line

up our own lives with the course of history itself and say ‘I was there’.” It could be considered rather Orwellian that these news memories are malleable.

Footnotes.

1. In order to keep the current analysis focused and clear, 96 participants who were randomly assigned to a third condition are not analyzed in the current study because by design they did not participate in the interview (see Patihis, 2012). The random assignment was set up on the computer to place one half of participants into the experimental condition, and one quarter into each of the two control conditions, resulting in 202 in the experimental condition and 95 in the human genome control condition in the main article.
2. By definition, the Human Genome condition ($n = 95$) reveals nothing related to false memories at the questionnaire stage, hence you only see those participants in the analysis of the interview that came 40 minutes later.
3. Because $n = 2$ we refrain from analyzing and discussing these individuals in depth in the article. We might speculate that in these individuals the initial suggestion had a delayed effect, although further research is required to establish this.

References

- Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences, 21*, 294–299.
- Ceci, S. J., Huffman, M. L. C., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition, 3*, 388–407.
- Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 42*, 304–320.
- Christianson, S. Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology, 1*, 225–239.
- Crombag, H. F. M., Wagenaar, W. A., & van Koppen, P. J. (1996). Crashing memories and the problem of ‘source monitoring’. *Applied Cognitive Psychology, 10*, 95–104.
- Garfinkel, S. N., Dienes, Z., & Duka, T. (2006). The effect of alcohol and repetition at encoding on implicit and explicit false memories. *Psychopharmacology, 188*, 498–508.
- Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin and Review, 3*, 208–214.
- Granhag, P. A., Stromwall, L., & Billings, J. (2003). I’ll never forget the sinking ferry: how social influence makes false memories surface. In M. Vanderhallen, G. Vervaeke, P. J. van Koppen, & J. Goethals (Eds.), *Much ado about crime: Chapters on psychology and law* (pp. 129–140). Politeia: Brussels.

- Gustavsson, J. P., Bergman, H., Edman, G., Ekselius, L., Von Knorring, L., & Linder, J. (2000). Swedish universities Scales of Personality (SSP): construction, internal consistency and normative data. *Acta Psychiatrica Scandinavica*, *102*, 217–225.
- Jelicic, M., Smeets, T., Candel, I., van Suijdam, M., & Merckelbach, H. (2006a). No, I don't remember seeing video footage of the killing of Theo van Gogh! Misinformation manipulations do not always elicit false memories. In K. Nixon (Ed.), *Forensic Recall and Eyewitness Testimony* (pp. 21–25). London, UK: IA-IP Publishing.
- Jelicic, M., Smeets, T., Peters, M. J. V., Candel, I., Horselenberg, R., & Merckelbach, H. (2006b). Assassination of a controversial politician: Remembering details from another non-existent film. *Applied Cognitive Psychology*, *20*, 591–596.
- 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.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517.
- Levine, L. J., Whalen, C. K., Henker, B., & Jamner, L. D. (2005). Looking back on September 11, 2001: Appraised impact and memory for emotions in adolescents and adults. *Journal of Adolescent Research*, *20*, 497–523.
- Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*, 720–725.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology*, *4*, 19–31.

- Lynn, S. J., Lilienfeld, S. O., Merckelbach, H., Giesbrecht, T., McNally, R. J., Loftus, E. F., ... & Malaktaris, A. (2014). The trauma model of dissociation: Inconvenient truths and stubborn fictions. Comment on Dalenberg et al.(2012). *Psychological Bulletin*, *140*, 896–910.
- Mazzoni, G. (2002). Naturally occurring and suggestion-dependent memory distortions: The convergence of disparate research traditions. *European Psychologist*, *7*, 17–30.
- McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, *114*, 1–16.
- Merckelbach, H. (2004). Telling a good story: Fantasy proneness and the quality of fabricated memories. *Personality & Individual Differences*, *37*, 1371–1382.
- Neisser, U. (1982). Snapshots or Benchmarks. In U. Neisser (ed.) *Memory Observed: Remembering in Natural Contexts*. Pp. 43-48. San Francisco, CA: Freeman/
- Ost, J., Granhag, P., Udell, J., Hjelmsäter, E. R. (2008). Familiarity breeds distortion: The effects of media exposure on false reports concerning media coverage of the terrorist attacks in London on 7 July 2005. *Memory*, *16*, 76–85.
- Ost, J., Hogbin, I., & Granhag, P. (2006). Altering false reports via confederate influence. *Social Influence*, *1*, 105–116.
- Ost, J., Vrij, A., Costall, A., & Bull, R. (2002). Crashing memories and reality monitoring: Distinguishing between perceptions, imaginings and false memories. *Applied Cognitive Psychology*, *16*, 125–134.
- Patihis, L. (2012). *Memory Distortions: Individual Differences and Paradigm Comparisons* (Master's thesis). University of California, Irvine. UMI 1512095.

- Patihis, L., Frenda, S. J., LePort, A. K., Petersen, N., Nichols, R. M., Stark, C. E., ... & Loftus, E. F. (2013). False memories in highly superior autobiographical memory individuals. *Proceedings of the National Academy of Sciences, 110*, 20947–20952.
- Patihis, L., & Younes Burton, H. J. (in press). False memories in therapy and hypnosis before 1980. *Psychology of Consciousness: Theory, Research, and Practice*.
- Porter, S., Birt, A. R., Yuille, J. C., & Lehman, D. R. (2000). Negotiating false memories: Interviewer and rememberer characteristics relate to memory distortion. *Psychological Science, 11*, 507–510.
- Porter, S., Taylor, K., & ten Brinke, L. (2008). Memory for media: Investigation of false memories for negatively and positively charged public events. *Memory, 16*, 658–666.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–814.
- Sheehan, P. W., & Tilden, J. (1983). Effects of suggestibility and hypnosis on accurate and distorted retrieval from memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*, 283–293.
- Sjödén, B., Granhag, P. A., Ost J., & Hjelmsäter, E. R. (2009). Is the truth in the details? Extended narratives help distinguishing false “memories” from false “reports.” *Scandinavian Journal of Psychology, 50*, 203–210.
- Smeets, T., Jelicic, M., Peters, M. J. V., Candel, I., Horselenberg, R., & Merckelbach, H. (2006). Of course I remember seeing that film! How ambiguous questions generate crashing memories. *Applied Cognitive Psychology, 20*, 779–789.

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.

Wilson, K., French, C.C. (2006). The relationship between susceptibility to false memories, dissociability, and paranormal belief and experience. *Personality and Individual Differences, 41*, 1493–1502.

Wright, D. B., & Loftus, E. F. (1999). Measuring dissociation: Comparison of alternative forms of the dissociative experiences scale. *American Journal of Psychology, 112*, 497–519.

Figure Captions

Figure 1. A simplified representation of the study design.

Figure 2. Percentages of participants by experimental condition who demonstrated a detailed Interview False Memory.

Figure 3. Percentage of participants in each age category who demonstrated a Persistent False Memory in the study. To convert the x-axis to the age of the participants at the time of the study, add 10 years.

Supplemental Material
Crashing Memory Computer Questionnaire Materials
United 93 Crash Footage Condition

Memories of News Events

Now we would like to gather some information about how well you remember news events. Please answer each of the following questions to the best of your ability.

1. As you may know, on September 11, 2001, United Airlines Flight 93 crashed in a field near Shanksville, Pennsylvania, killing all 44 people on board. Video footage of the plane crashing, taken by one of the witnesses on the ground, has been well publicized both by the news media and on the internet. Have you seen the video? (Check one)

Yes

No

2. How well can you remember having seen the video?

1

2

3

4

5

6

7

8

9

10

(No memory at all)

(Very clear memory)

Please answer the following questions about your memory of the video footage of the United 93 crash.

3. The plane

- (a) came down vertically, nose down and almost without forward speed
- (b) slid into the ground almost horizontally and at considerable speed
- (c) I can't remember

4. After the impact

- (a) parts of the plane were visible in the wreckage
- (b) the plane's body disintegrated
- (c) the fire and smoke made it impossible to tell
- (d) I can't remember

5. The video footage was

- (a) very clear, you can see and hear exactly what is happening
- (b) fuzzy, it is difficult to tell what is happening
- (c) I can't remember

6. The footage was

- (a) less than 60 seconds long
- (b) between 1 and 2 minutes long
- (c) longer than 2 minutes
- (d) I can't remember.

[Memory Characteristic Questionnaire]

MCQ1. My memory for the footage is

1

2

3

4

5

6

7

I can't remember

(dim)

(sharp/clear)

MCQ2. My memory for the footage involves visual detail

1 2 3 4 5 6 7 I can't remember
(little or none) (a lot)

MCQ3. My memory for the footage involves sound

1 2 3 4 5 6 7 I can't remember
(little or none) (a lot)

MCQ4. Overall vividness of my memory of the footage is

1 2 3 4 5 6 7 I can't remember
(vague) (very vivid)

MCQ5. My memory for the location where the footage takes place is

1 2 3 4 5 6 7 I can't remember
(vague) (very vivid)

MCQ6. The video footage seems

1 2 3 4 5 6 7 I can't remember
(short) (long)

MCQ7. The overall tone of my memory for the video footage is

1 2 3 4 5 6 7 I can't remember
(extremely negative) (extremely positive)

MCQ8. I remember how I felt at the time I first saw the footage

1 2 3 4 5 6 7
(not at all) (definitely)

MCQ9. Feelings at the time that I first saw the video were

1 2 3 4 5 6 7 I can't remember
(extremely negative) (extremely positive)

MCQ10. Feelings at the time that I first saw the video were

1 2 3 4 5 6 7 I can't remember
(not intense) (very intense)

MCQ11. As I am remembering the footage now, my feelings are

1 2 3 4 5 6 7 I can't remember
(not intense) (very intense)

MCQ12. Since I saw the footage, I have thought about it

1 2 3 4 5 6 7
(not at all) (many times)

MCQ13. Since I saw the footage, I have talked about it

1 2 3 4 5 6 7
(not at all) (many times)

Control Condition (Human Genome)

Memories of News Events

Now we would like to gather some information about how well you remember news events. Please answer each of the following questions to the best of your ability.

1. On February 1, 2001, the Human Genome Project international consortium announced the publication of a draft sequence and initial analysis of the human genome the genetic blueprint for a human being. The pioneering paper appeared in the February 15 issue of the journal Nature. Did you hear or see this news event? (Check one)

Yes

No

Please answer the following questions about your memory of news for the Human Genome Project.

- 3. The Human Genome Project
 - (a) was complete in 2001
 - (b) was fully completed in 2004
 - (c) I can't remember

- 4. After the news of the Human Genome Project
 - (a) It was widely reported in all media
 - (b) It was reported only in newspapers
 - (c) It was reported on TV news only
 - (d) I can't remember

- 5. The news coverage of the event was
 - (a) very clearly communicated
 - (b) not clearly communicated
 - (c) I can't remember

- 6. The first news report was
 - (a) less than 60 seconds long
 - (b) between 1 and 2 minutes long
 - (c) longer than 2 minutes
 - (d) I can't remember.

[Memory Characteristic Questionnaire]

MCQ1. My memory for the initial news report is

1	2	3	4	5	6	7	I can't remember
(dim)						(sharp/clear)	

MCQ2. My memory for the initial news report involves visual detail

1	2	3	4	5	6	7	I can't remember
(little or none)						(a lot)	

MCQ3. My memory for the initial news report involves sound
 1 2 3 4 5 6 7 I can't remember
 (little or none) (a lot)

MCQ4. Overall vividness of my memory of the initial news report is
 1 2 3 4 5 6 7 I can't remember
 (vague) (very vivid)

MCQ5. My memory for the location where the initial news report takes place is
 1 2 3 4 5 6 7 I can't remember
 (vague) (very vivid)

MCQ6. The initial news report seemed
 1 2 3 4 5 6 7 I can't remember
 (short) (long)

MCQ7. The overall tone of my memory for the initial news report is
 1 2 3 4 5 6 7 I can't remember
 (extremely negative) (extremely positive)

MCQ8. I remember how I felt at the time I first saw the initial news report
 1 2 3 4 5 6 7
 (not at all) (definitely)

MCQ9. Feelings at the time that I first learned about the initial news report were
 1 2 3 4 5 6 7 I can't remember
 (extremely negative) (extremely positive)

MCQ10. Feelings at the time that I first learned about the initial news report were
 1 2 3 4 5 6 7 I can't remember
 (not intense) (very intense)

MCQ11. As I am remembering the initial news report now, my feelings are
 1 2 3 4 5 6 7 I can't remember
 (not intense) (very intense)

MCQ12. Since I saw the initial news report, I have thought about it
 1 2 3 4 5 6 7
 (not at all) (many times)

MCQ13. Since I saw the initial news report, I have talked about it
 1 2 3 4 5 6 7
 (not at all) (many times)

Interview Script

The last thing is a 5 minute recorded interview that is anonymous and confidential, so we will just use your participant number during the interview, and not your name. Please speak as freely as possible and answer all of questions with as much detail as you can. We are looking for what you really remember, there are no right or wrong answers, and it is okay if there is anything you cannot remember.

[AT THE BEGINNING OF EACH INTERVIEW START THE RECORDING WITH THE FOLLOWING WORDS]

“This is participant number _ _ _ _”

I want to ask you a few questions about how well you remember news events. As you might recall, on September 11, 2001, two planes were flown into the world trade center in New York City, one plane was flown into the Pentagon in Washington DC, and another plane, United 93 crashed into a field in rural Pennsylvania. The plane crash in Pennsylvania is the event we are interested in asking you about.

The other crashes on 911 have already been studied, so we are focusing only on United 93, the one that crashed in a field in Pennsylvania.

Are you familiar with this event?

Can you tell me what you remember about the event?

[LISTEN TO CHECK THEY UNDERSTAND IT IS ABOUT THE CRASH INTO THE FIELD]

As you might know, a witness on the ground in Pennsylvania took some video of the plane crashing and it has been widely shown on TV news and the Internet in the months and years since the attack.

Do you remember seeing that footage?

[IF YES]

Can you tell me what you remember about the footage?

[FREE RESPONSE]

Can you describe how the plane moved in the footage?

Do you remember how the plane crashed in the video?

How did you feel when you saw the footage of United 93 crash in Pennsylvania?

How vivid is your memory of that footage of the crash?

Do you remember how long the video is?

Do you remember if the video had sound?

If you did see the footage, where did you see it first? (Was it on the internet or TV, if so which channel)

Can you remember any additional details? Take a moment to think if you like.

Okay, now I'd like you to tell me how well you can remember having seen the video on a scale from 1 to 10 , where 1 means no memory at all and 10 means a very clear memory. **[END OF INTERVIEW]**

[IF NO]

Sometimes memories fade so we can't remember them, especially ones that are unpleasant or traumatic. However, we can use techniques that can help us find those memories.

[TALK SLOWLY, RAPPORT] If you don't mind, I'd like for you to close your eyes for a few moments. I would like you to use your imagination and try to picture what the footage may have looked like. Imagine you are watching it on your television or your computer screen. You are watching a video of the plane crashing, taken by a witness who is standing in a grassy field near the crash site. Just take a few moments and let any images or sounds come into your head.

[LET MORE THAN 30 SECONDS PASS DURING THIS EXERCISE, INCLUDING 10 SECONDS OF SILENCE AFTER THE LAST SENTENCE]

Keeping your eyes closed, can you describe to me what you are seeing in your minds eye?

[PARTICIPANT RESPONDS – WAIT FOR THEM TO STOP TALKING AND WAIT A FEW SECONDS BEFORE MOVING ON]

Can you describe how the plane moves?

Describe how the plane crashes in the video?

What does the aftermath look like?

What about the people filming the video, do you hear them talking?

(You can now open your eyes).

Actually, several of the details you are giving me are exactly consistent with the video. So that's really good. Do you feel like you might be remembering the footage?

Do you remember how long the video is?

Where would you have been when you first saw it, right after 9/11?

Do you remember how you felt after seeing it?

Can you remember any additional details? Take a moment to think, if you like.

Okay, now I'd like you to tell me how well you can remember having seen the video on a scale from 1 to 10, where 1 means no memory at all and 10 means a very clear memory.

Okay, now that the interview is over, I would like to ask you just one last question – Did you indicate that you had seen the United 93 footage in the computer questionnaire?

[If Yes]: Can you tell me why your answer changed from yes to no between the computer questionnaire and this interview? **[END OF INTERVIEW]**

Supplemental Results

Relationship between False Memories in Interview and Interviewer Characteristics

Because the interview was a face to face exchange, we wondered if the interviewer's gender, age, socioeconomic status, empathy, or personality traits might be related to Interview False Memory rates. The interviewers were 19 research assistants (5 male, 14 female; ages 21-47). To ensure a fair comparison, we examined this within the Experimental condition. We found little to no evidence that interviewer characteristics mattered, with the only finding close to being marginal was for the personality measure verbal trait aggression: False Memory production in interviews involved interviewers with slightly higher verbal trait aggression ($M = 2.71$, $SD = .29$) than those without a False Memory ($M = 2.63$, $SD = .31$), $t(200) = 1.49$, $p = .138$. Apart from that, False Memory production in the interview was not associated with interviewer characteristics, $ps > .293$. The interviewer personality traits included somatic anxiety, psychic anxiety, stress susceptibility, impulsivity, adventure seeking, detachment, social desirability, embitterment, and physical trait aggression (subscales of SSP).

Excerpts from Interview False Memories

Here, we document some of the wording used by participants in the interview who demonstrated an Interview False Memory.:

Interviewer (I): Okay, so can you tell me what you do remember about the footage?

Participant 1007 (P): I just remember a little like, the plane falling past trees and that's all I remember seeing.

I: Okay, and can you tell me, or can you describe how the plane moved in the footage?

P: It was going down kind of at an angle.

I: Like what kind of, like, degree of angle would you say?

P: Like 45 degrees. ...

I: Alright, and how did you feel when you saw the footage of United 93 crash in Pennsylvania?

P: Kind of upset, and a little angry.

I: Okay, and how vivid is your memory of the footage of the crash?

P: Not to vivid actually, I was kind of young.

I: Okay, and do you remember how long the video is?

P: Maybe like a minute or so, two minutes.

I: And do you remember if the video had any sounds?

P: Not really I think there might have been a guy or someone in the background talking over it but I don't remember it.

I: Okay, so if you saw the footage where did you see it first, was it on the internet, on the TV?

P: TV

I: On what channel, do you, do you remember?

P: Probably FOX, channel 11.

I: Okay, and can you remember any additional details? You can take a moment to think anything you'd like.

P: No, I just remember firefighter crews trying to put out the flames, and then umm, that's about it.

I: Okay, now I would like for you to tell me how well you remember having seen the video on a scale from 1 to 10. So, 1 would mean no memory at all and 10 would stand for a very clear, vivid memory.

P: Probably a 3 or 4.

In the above example, the fragility of the newly forming false memory is demonstrated in the tentative language (e.g. “probably”). Nevertheless we see evidence of emerging false details that could become permanently consolidated with time and reinforcement (e.g. “the plane falling past trees” and “45 degrees”). This combination of false details but with some degree of uncertainty characterized many of the participants who demonstrated a false memory in the interview. Another example of such a labile emerging false memory is given below:

I: ...Do you remember seeing the footage?

Participant 1044 (P): Yeah I saw it on the news / TV.

I: Can you tell me what you remember about the footage?

P: Oh, it just burst into flames when it crashed into the floor...

I: Do you remember how the plane crashed in the video?

P: I just remember it hit the floor. ...

I: How did you feel when you saw the footage of the United 93 crash in Pennsylvania?

P: Kind of like shocked and sad for the people.

I: How vivid is your memory of the footage of the crash?

P: Okay I guess.

I: Do you remember how long the video is?

P: Just a couple of seconds. ...

I: Can you remember any additional details about United 93.

P: I think they followed up with saying that it crashed into the floor because some guys tried to distract the hijackers so it didn't crash into the Whitehouse. So that's why it landed on the floor.

I: Now I would like for you to tell me how well you remember having seen the video on a scale from 1 to 10. So, 1 would mean no memory at all and 10 would stand for a very clear, vivid memory.

P: I give it like a 5.

Results for Young Adults (ages 18–22) Only

In this analysis, we exclude participants from age 23 to 51 ($n = 22$) for two reasons. First, the majority of our participants were between 18 and 22 ($n = 271$) and we wanted here to focus on participants between the ages of 18 and 22 because those were in mid-childhood (ages 8 to 12) when the target events of 9/11 occurred. In this young adult sub-sample the mean age was 19.4 ($SD = 1.2$), with 29% at age 18, 33% at age 19, 17% at age 21, and 5% at age 22.

Computer Questionnaire

Of the 183 18–22 year olds assigned to the United 93 Crash Footage condition, when asked "have you seen the video?" 38.8% (70) indicated "yes," they had seen the United 93 crash footage. Of the 38.8% who indicated "yes," they had seen the footage, 91.4% gave at least one false detail (how the plane moved, clarity of footage, or length of footage). Specifically, of those 36.6% who said "yes," 62.0% gave a detail about the plane moving, 59.2% gave a detail about after the impact, 67.6% gave some detail about the clarity of the footage, and 49.3% gave a detail about the length of the footage (see Supplemental Material for question wording on the 4 detail questions). Of *all* the 183 young adult participants, 43.4% indicated some detail of the nonexistent plane crash footage.

When asked "how well can you remember having seen the video?" 57.7% indicated 1 (*no memory at all*) on a scale from 1 to 10, and 42.3% indicated a 2 or above. 10.2% indicated a score of 5 or above, with one participant (.5%) indicating 10 (*a very clear memory*). Of the 70

that had indicated “yes” they had seen the footage, all (70; 100%) indicated a score of 2 or above on this scale, with 25.7% (18) indicating a score of 5 or above. In contrast, of the 112 who indicated “no” they had not seen the video, 93.8% (105) indicated a score of 1 (*no memory at all*) on this 1–10 scale.

False Belief versus False Memory. Using the criteria described in the main text, 64 (35.0%) of the 183 young adults indicated a detailed Questionnaire False Memory, 6 (3.3%) indicated a false belief, and 113 (61.7%) were correct.

Audio Recorded Interview

About 40 minutes after the computer questionnaire, those in the United 93 Crash footage (Experimental) condition, and the Human Genome (Control) condition participated in the face to face recorded interview with a research assistant. Of the 271 young adults who participated in the crashing memory interview, 44 (16.2%) said they had seen the footage, 20 (7.4%) said unsure/maybe, and 207 (76.4%) said they had not seen crash footage of United 93. This figure of 16.2% saying “yes” in the interview compares to 38.8% that indicated “yes” in the forced-choice computer questionnaire earlier. In the interview, when asked how well they remembered having seen the video on a scale from 1 to 10, 37.3% (101 out of 271) indicated 1 (*no memory at all*), 62.3% indicated a 2 or above, with 16.2% (44) indicated a score of 5 or above.

False Memory versus False Belief. Using criteria mentioned in the main article, we identified 28 (10.3%) as having an Interview False Memory, 16 (5.9%) having a false belief, and 227 (83.8%) showing neither.

Persistent False Memory. By the metric defined in the main article, 24 (13.1%) young adults exhibited a Persistent False Memory in questionnaire and interview (out of the 183 who participated in the suggestive questionnaire *and* interview).

Comparing those Familiar to those Not Familiar with the Specific News Story

Although all participants were familiar with the 9/11 general news story, only 127 (67.4%) of the 183 young adults in the Experimental condition were specifically familiar with the United 93 story. In the *questionnaire*, the 127 participants familiar with the United 93 news story, compared to those 60 not familiar, were marginally more likely to indicate a Questionnaire False Memory (39.4% vs. 25.0%; $\chi^2(1, N = 183) = 3.53$, exact $p = .066$, Cramer's $V = .139$). In the *interview* those familiar with the United 93 news story, compared to those not familiar, were more likely to indicate an Interview False Memory (15.4% vs. 1.0%; $\chi^2(1, N = 271) = 13.9$, exact $p < .001$, Cramer's $V = .256$). Similarly, those familiar with the United 93 news story, compared to those not familiar, were more likely to indicate a Persistent False Memory throughout the questionnaire and interview (18.9% vs. 0.0%; $\chi^2(1, N = 183) = 12.2$, exact $p < .001$, Cramer's $V = .258$).

Comparing Experimental to Control Group

Effect of Prior Suggestion on Interview False Memory. Some of the young adults who participated in the audio recorded interview had gone through a suggestive questionnaire about United 93 on the computer 40 minutes earlier (Experimental condition), while others had been randomly assigned to a non-suggestive neutral questionnaire about the 2001 human genome news story (Control condition). A higher proportion of those in the Experimental condition exhibited Interview False Memories (14.2%; 26 of 183) compared to those in the Control condition (2.3%, 2 of 88), $\chi^2(1, N = 271) = 9.14$, exact $p = .002$, Cramer's $V = .184$. This six-fold difference in False Memory proportions in the interview shows the potent effect of the suggestive questionnaire 40 minutes earlier.

Comparing those with Persistent False Memory vs. those Without

Demographics. There was no difference between those who had a Persistent False Memory on gender, age, or on whether the participant was a psychology-related major or not ($p > .735$). Those identifying their ethnicity as Middle Eastern had significantly higher rates of Persistent False Memory (46.2%; 6 of 13) compared to those of other ethnicities (10.6%; 18 of 170), $\chi^2(1, N = 183) = 13.41$, exact $p = .003$, Cramer's $V = .271$. This association between Middle Eastern ethnicity and Persistent False Memory was not effected by familiarity with the United 93, and was only slightly partially mediated by high-arousal negative emotion in the week following 9/11 (logistic regression: before mediation: Middle Eastern ethnicity $\beta = 1.98$, $p = .001$; after mediation: $\beta = 1.74$, $p = .006$)

Reported Emotion in Week After 9/11. Those who demonstrated a Persistent False Memory reported higher ratings of how often they had felt negative emotions in week after following 9/11 ($M = 65.0$, $SD = 28.4$), compared to those not exhibiting a Persistent False Memory ($M = 48.5$, $SD = 27.1$), $t(181) = 2.76$, $p = .006$. This difference was found regardless of the types of negative emotions, including high arousal, pre-goal, and post-goal negative emotions ($p < .039$). These emotions in the week after 9/11 were reported approximately 10 years after 9/11.

Alcohol Use. Those young adults who demonstrated a Persistent False Memory reported higher frequency of consuming alcohol than did those who did not show a Persistent False Memory ($M = 2.79$: just below the anchor *twice a month*, $SD = 1.56$), compared to those not exhibiting a Persistent False Memory ($M = 2.07$, just above the anchor *once a month*, $SD = 1.31$), $t(181) = 2.46$, $p = .015$. This Likert scale was a fully anchored scale from $1 = \textit{never}$ to $9 = \textit{everyday}$ (from LaBrie, Hummer, Grant, & Luc, 2010). A summed composite of this and three

other alcohol use questions revealed a marginally statistically significant difference, $t(181) = 1.82, p = .071$.

Fantasy Proneness and Dissociation. Those with a Persistent False Memory were significantly higher of fantasy proneness ($M = 11.4, SD = 4.31$; Creative Experiences Questionnaire) compared to those with no Persistent False Memory ($M = 9.32, SD = 4.07$), $t(181) = 2.29, p = .023$. There were no differences between those with Persistent False Memory and those without on dissociation ($p = .860$; DES-C).

Relationship between Interview False Memories and Interviewer Characteristics

We found little to no evidence that interviewer characteristics (gender, age, SES, personality) mattered: False Memory production in the interview was not associated with interviewer characteristics, $ps > .196$.

Summary. The memory distortion experiment successfully implanted false memories for an upsetting event that occurred 10 years earlier, in the participant's childhood (ages 8–12). It reliably produced false memory reports of seeing the nonexistent footage of the United 93 crash in a sizable minority of subjects in both the computer questionnaire, and in our more rigorous follow-up interview. In that interview, we took pains to be clear about what crash we were referring to, reiterated it was a crash into a *field*, we eliminated any forced-choice, and we asked follow up questions to help us distinguish between a false memory and a false belief. We were conservative in that we required three prerequisite necessary conditions (a “yes”, a score not including 1 = *no memory at all* on a 1–10 scale, and a false detail of the actual crash) when distinguishing False Memory from false beliefs in both the questionnaire and the interview. Even with the necessary condition of a false detail accompanying the false report, we still found 13.1% had a Persistent False Memory in both the questionnaire and the interview. Of those familiar

with the United 93 news story, 18.9% had a Persistent False Memory. Prior suggestion 40 minutes earlier in the questionnaire had an additive effect on False Memory rates in the interview, with approximately a six-fold difference in percentages. Middle Eastern participants had higher rates of Persistent False Memory than people of other ethnicities. Frequent negative emotions in the week after 9/11, alcohol use, and fantasy proneness were associated with higher Persistent False Memory.

Table S1

Pearson r effect sizes of associations between items on the MCQ relating to memory of the footage in the questionnaire, and a subsequent Interview False Memory occurring 40 minutes later (N = 202).

	<i>r</i>	<i>p</i>
MCQ1 My memory for the footage is <i>Anchors: 1 = dim, 7 = sharp/clear</i>	.311	< .001
MCQ2 My memory for the footage involves visual detail <i>1 = little or none, 7 = a lot</i>	.339	< .001
MCQ3 My memory for the footage involves sound <i>1 = little or none, 7 = a lot</i>	.310	< .001
MCQ4 Overall vividness of my memory of the footage is <i>1 = vague, 7 = very vivid</i>	.328	< .001
MCQ5 My memory for location where the footage takes place is <i>1 = vague, 7 = very vivid</i>	.374	< .001
MCQ6 The video footage seems <i>1 = short, 7 = long</i>	.201	.004
MCQ7 The overall tone of my memory for the video footage is <i>1 = extremely negative, 7 = extremely positive</i>	.318	< .001
MCQ8 I remember how I felt at the time I first saw the footage <i>1 = not at all, 2 = definitely</i>	.143	.043
MCQ9 Feelings at the time that I first saw the video were <i>1 = extremely negative, 7 = extremely positive</i>	.348	< .001
MCQ10 Feelings at the time that I first saw the video were <i>1 = not intense, 7 = very intense</i>	.409	< .001
MCQ11 As I am remembering the footage now, my feelings are <i>1 = not intense, 7 = very intense</i>	.314	< .001
MCQ12 Since I saw the footage, I have thought about it <i>1 = not at all, 7 = many times</i>	.131	.063
MCQ13 Since I saw the footage, I have talked about it <i>1 = not at all, 7 = many times</i>	.135	.055

Note. Each question had a Likert scale from 1 to 7, with an option “I don’t remember” coded as 0. Anchors are shown in the table below each item. Intercorrelations between MCQ questions were between $r = .3$ and $.9$, indicating collinearity that would preclude a logistic regression model.