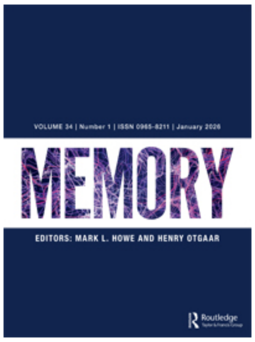

Memory

ISSN: 0965-8211 (Print) 1464-0686 (Online) Journal homepage: www.tandfonline.com/journals/pmem20

Trained but still tricked: source sensitisation training fails to reduce false memory reports

Charlotte A. Bücken, Paul Riesthuis, Giorgia Caon, Alexandra Cucu & Henry Otgaar

To cite this article: Charlotte A. Bücken, Paul Riesthuis, Giorgia Caon, Alexandra Cucu & Henry Otgaar (2026) Trained but still tricked: source sensitisation training fails to reduce false memory reports, Memory, 34:1, 19-33, DOI: 10.1080/09658211.2025.2585095

To link to this article: https://doi.org/10.1080/09658211.2025.2585095

MEMORY
2026, VOL. 34, NO. 1, 19–33
https://doi.org/10.1080/09658211.2025.2585095

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Trained but still tricked: source sensitisation training fails to reduce false memory reports

Charlotte A. Bücken [a,b], Paul Riesthuis [a,b], Giorgia Caon[a,c], Alexandra Cucu[a] and Henry Otgaar[a,b]

[a]Leuven Institute of Criminology, Faculty of Law and Criminology, KU Leuven, Leuven, Belgium; [b]Forensic Psychology Section, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands; [c]Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

ABSTRACT
False autobiographical memories can have serious implications in legal settings, where the case outcomes may hinge entirely on memory-based eyewitness testimony. This study investigated whether a sensitisation memory training could reduce false autobiographical memory reports. We employed a blind implantation method in which participants ($N = 294$) indicated whether various childhood events had happened to them. Participants were then told they had confirmed five events – one of which was false – and were asked to rate their memory and belief. In session two, 15% (44/294) of participants reported a false belief and an additional 3.4% (10/294) a false memory, meaning that a total of 18.4% made a false report. Before session three, participants were randomly assigned to receive either the memory training or a distractor task, then repeated the false memory procedure. Contrary to our expectations, the training did not reduce false reports. Instead, false beliefs (SMT: 20.4%, 28/137, Control: 22.3%, 31/139) and false memories (SMT: 5.1%, 7/137, Control: 2.9% 4/139) increased in session three. The findings suggest false memories elicited in the blind implantation paradigm might be particularly resistant to correction.

ARTICLE HISTORY
Received 16 July 2025
Accepted 28 October 2025

KEYWORDS
Blind implantation; false autobiographical memory; false memory implantation; false memory reversal; source sensitization memory training

In legal settings, eyewitness memory accounts can have tremendous importance for legal decision-making. This is especially the case for crimes such as child abuse, where eyewitness testimony might be the only available evidence (Leander, 2010; Lyon et al., 2020), as other forensic evidence is frequently absent (Lamb et al., 2011; but see also Herman, 2010). In such cases, legal decisions of whether a suspect is guilty or not may rest entirely on testimonial evidence. Therefore, it is important that eyewitness accounts are valid, accurately reflecting events as they happened (Orbach et al., 2000; Vredeveldt et al., 2022). If they do not accurately reflect what happened, this can lead to false testimony and can carry grave consequences, such as wrongful convictions (www.innocencenetwork.org). An example of such a case is that of Gregory Kelly who was wrongfully convicted in 2014 for super aggravated sexual assault of a four year old boy at an at-home day care centre (https://exonerationregistry.org/cases/12735). After several leading interviews, the alleged victim accused Kelly of sexual assault, and the latter was convicted and received a 25 year long sentence. The only evidence linking Kelly to the alleged crime was the testimony of the boy – in fact, exculpatory evidence was ignored in the investigation. The conviction was overturned in 2019 on the bais of "actual innocence" after someone else confessed to the crime.

Several factors such as being asked leading questions, drinking alcohol, or even lying can undermine eyewitness memory by contributing to omission errors (e.g., forgetting details of an event) or commission errors (e.g., misremembering details about experienced events; e.g., Kloft et al., 2021). However, previous research also demonstrated that people can falsely remember entire autobiographical events (Loftus & Pickrell, 1995; Murphy et al., 2023; for a review: Scoboria et al., 2017). An often-used method to induce such "rich false memories" is the implantation method (Scoboria et al., 2017). As the name suggests, in this method experimenters attempt to "implant" a false memory in the mind of the participants. To do so, the researchers contact the participants' parents to obtain information about several events that happened in the

CONTACT Charlotte A. Bücken charlotte.buecken@kuleuven.be Leuven Institute of Criminology, Faculty of Law and Criminology, KU Leuven, Herbert Hooverplein 9, Leuven 3000, Belgium; Forensic Psychology Section, Faculty of Psychology and Neuroscience, Maastricht University, Universiteitssingel 40, Maastricht 6229ER, The Netherlands

participant's childhood. More importantly, they also confirm with parents that the to be implanted event (e.g., being lost in a shopping mall; Loftus & Pickrell, 1995) has not been experienced. Then, participants are asked to report what they remember about these events their parents told the researchers about (i.e., some are truly experienced events – according to the parents – and one is the non-experienced false event). A mega-analysis of such implantation studies found that in laboratory settings around 30% of participants produce false memories, while others gain false beliefs (23%; Scoboria et al., 2017). Certain factors can increase such implantation or false memory rates, and depending on suggestibility conditions the number of participants who produce a false memory goes up to 48% (Scoboria et al., 2017). For example, simply asking someone to imagine a childhood event repeatedly and going through imagination exercises can inflate confidence that this event really happened, also called "imagination inflation" (Garry et al., 1996), and thus contribute to false memory rates.

In the false memory literature there has sometimes been dispute regarding whether "implanted" false memories are actually false *memories* or false *beliefs* (Goodman et al., 2025; Scoboria et al., 2017). Although the concepts of (false) memories and (false) beliefs are related, they are not the same. Specifically, *memory* refers to the vivid imagery one experiences when thinking about an experienced event whereas *belief* refers to the truth status someone attaches to a certain experience – for example, whether they believe that they truly experienced it (Scoboria et al., 2004). Both can have tremendous consequences in legal settings as they can affect behaviour (e.g., Bernstein & Loftus, 2009; Geraerts et al., 2008), for example by affecting eyewitness misidentifications (i.e., such as in the case of Gregory Kelley). Yet, when someone creates a false *memory* and perhaps even reports more false details, this can lead to additional false testimony that might add false incriminating evidence. To combat the adverse consequences of false memories and beliefs, we will examine whether they can be reversed so as to reduce their potential damage.

Source monitoring framework

There are different theoretical accounts as to why we might misremember details or entire events that never happened to us. A dominant theory is the Source Monitoring Framework (SMF; Johnson et al., 1993; Johnson & Raye, 1981). According to the SMF ourmemories can have different origins also called sources which specify how the event was encountered and entered into memory. Indeed, we can have memories stemming from internal sources, such as imagining or dreaming that an event happened, or external sources such as perception. "Source monitoring" then refers to the automatic process in which we differentiate between memories that stem from different sources (Johnson et al., 1993; Johnson & Raye, 1981). This process is based in heuristics; for example, memories from external sources tend to hold more perceptual and contextual details. Memories from internal sources, however, tend to contain more cognitive operations. Thus, when a memory contains many perceptual details, we often heuristically attribute it to a source in external experience.

However, because source monitoring is based in such heuristic judgements, errors can occur. For example, when a memory from an internal source (e.g., a dream or imagination exercise) contains a lot of perceptual characteristics it might be wrongfully attributed to an external source. Indeed, studies show that simply imagining an action repeatedly can later on lead to a false belief in participants that they actually performed the action (Goff & Roediger, 1998; Nichols & Loftus, 2019). Another possible source monitoring error is attributing an event that someone else told us about (e.g., an interviewer who is telling us that our parents claimed this event happened to us) to an origin in real experience. Thus, non-experienced details or entire events can turn into false beliefs or memories when a source monitoring error occurs (Johnson et al., 1993). Improving the source monitoring process and judgements should be able to reduce the number of source monitoring mistakes that a person makes, and thus reduce the likelihood of reporting a false memory or belief.

Reversing false memories and beliefs

Indeed, some studies have demonstrated that false memories or beliefs could be reduced or even reversed (e.g., Blank & Launay, 2014; Oeberst et al., 2021). For instance, telling participants that they were exposed to false information or suggestion (i.e., "post-warnings") – can successfully reduce their susceptibility to making a false report later on by around 50% (Blank & Launay, 2014). Moreover, careful debriefing can also lead to retractions of rich false memories in participants, showing that these can in principle be reversed after implantation (Greene et al., 2024). However, directly telling someone that they received misinformation about a certain event is not applicable to a forensic context (e.g., when a witness or alleged victim is being interviewed or gives testimony). The reason is because in such contexts ground truth is oftentimes unknown and it is uncertain to what extent a witness was exposed to suggestions or false information at an earlier time point. Hence, some recent studies have examined other strategies aimed at reducing or reversing false memories (e.g., Oeberst et al., 2021; Spearing et al., 2025; Szpitalak et al., 2021) such as indirectly cueing participants to think about the sources of their memories (e.g., if they remember something because they truly experienced it, or because someone else has suggested it to them). One such indirect cue is informing participants about how memory works, and that errors can be made during remembering (Oeberst et al., 2021), also known as a sensitisation strategy, as it is meant to *sensitise* participants to certain issues in remembering.

To examine the effectiveness of sensitisation strategies on reversing false memories, Oeberst and colleagues (2021) attempted to implant an unexperienced event in participants' memory across three interviews. Over these interviews, between 27% (low suggestive techniques) and 56% (very suggestive techniques) of participants formed false memories or beliefs of plausible but non-experienced events (these differed per participant and included getting lost, running away, or being involved in a car accident, for example). Specifically, participants went from on average having no memory of the event to accepting that the event happened (i.e., a false belief) after the first and second interview and on average had at least a partial false memory after the third interview.[1] Then, during follow-up sessions, participants were presented with sensitisation techniques: in the fourth session, participants received a source sensitisation instruction, alerting them to possible external sources of memories, followed by another interview. In the fifth session, participants were alerted of the possibility of having false memories – a false memory sensitisation, also followed by an interview. At the end of the fifth session and after exposure to both sensitisation instructions, false memory rates went back to session 1 baseline (i.e., 15-25%) and participants on average indicated having "no memory" of the event (i.e., neither a false belief nor a (partial) false memory at the last follow-up). Thus, such sensitisation instructions, which could be used in forensic contexts without knowing the ground truth – were quite effective in reducing false reports (i.e., both false memories and false beleifs) in the Oeberst and colleagues (2021) study.

In another study, Spearing and colleagues (2025) showed that source sensitisation can reduce false memories when using the misinformation paradigm. The misinformation paradigm is a procedure in which participants are exposed to a stimulus (e.g., perpetrator holding a knife), receive misleading post event information (i.e., misinformation; e.g., the perpetrator held a gun) about it, and then have their memory tested. The canonical finding in this paradigm is that around 30% of participants report the misinformation when their memory for the initial stimulus is tested (Loftus & Klemfuss, 2023). Using this paradigm, Spearing and colleagues (2025) conducted two experiments in which participants were exposed to misleading information about a mock crime. In a next step, half of the participants received sensitisation instructions (a combination of source and false memory sensitisation, delivered via an educational video), whereas the other half watched an unrelated Youtube video on musical instruments. Finally, all participants completed a memory test. The misinformation effect was smaller for participants who had seen the sensitisation video compared to those who had not: Participants in the sensitisation condition were 64% correct about misled items, and non-sensitised participants around 59% (Experiment 1; percentages were similar in experiment 2: around 62% in the sensitised and 56% in the non-sensitised condition; Spearing et al., 2025). Thus, Spearing et al. (2025) showed that sensitisation instructions also successfully reduce false memory reports in misinformation paradigms.

Relatedly, Szpitalak and colleagues (2021) also tested the effects of a memory training on the formation false memories in a misinformation paradigm. They used a video of a theft as the stimulus and exposed participants to several misleading details in a description of the film. Subsequently, half of the participants received memory training, whereas the control condition received a training on social attitudes. This memory "training" was passive in nature (i.e., did not include an interactive part) and comprised watching a three minute film explaining different memory errors. Thereby, it was similar to the false memory sensitisation instructions used by Spearing and colleagues (2025). Szpitalak and colleagues (2021) found that their training was effective in increasing correct responding (i.e., rejecting misinformation and correctly identifying true information) of misinformed participants by on average one correct answer more than misinformed, untrained participants.

In summary, these studies have shown that sensitisation instructions, which directly teach participants about the different sources of memories (i.e., source sensitisation) and/ or about the existence of false memories (i.e., false memory sensitisation), can significantly reduce false memory reports in 1) an implantation paradigm (Oeberst et al., 2021) but also replicate in 2) a misinformation paradigm (Spearing et al., 2025; Szpitalak et al., 2021).

Effects of reversal strategies on true memory

Sometimes, strategies that are employed to reduce the reporting of false memories can have a negative side effect: warnings, for example, may lead to a so-called *tainted truth* effect, in which non-misled participants who receive a warning against misinformation "overcorrect" their memories for true experiences and decrease their true memory reports (Echterhoff et al., 2007; Szpitalak et al., 2021). Indeed, in the first experiment of the Szpitalak et al. (2021) study, participants were presented with a warning as a misinformation reduction strategy. When this warning was employed, the researchers found evidence of a tainted truth effect: participants who were not misled but received a warning answered correctly on less items compared to participants who were not misled and did not receive a warning. However, when testing for such a tainted truth effect in their second experiment, in which participants received a source sensitisation training, no such tainted truth effect arose (Szpitalak et al., 2021). In line with this, the other two studies that examined the effects of sensitisation strategies on true and false memory reports have not found any evidence that true memories were (negatively) affected by such instructions (Oeberst et al., 2021; Spearing et al., 2025).

Gaps in research

Oeberst and colleagues (2021) were the only ones who examined whether sensitisation instructions can reduce false reporting of *autobiographical,* rich false memories by using an implantation method – the other two studies examined memories for videos. However, one common criticism of the kind of implantation paradigm used by Oeberst and colleagues (2021) is that it relies entirely on parents' memory to clarify the ground truth of whether an event has or has not happened to the participant (Otgaar et al., 2022). Yet, parents might have forgotten themselves that certain events happened to their child. Thus, some false memories and beliefs reported in implantation studies might actually be true memories, which arguably might be more difficult to reduce than false memories. A way to circumvent the issue of involving parents' memory in the implantation process of autobiographical memories is by using a blind implantation paradigm (Otgaar et al., 2022) to examine the effectiveness of source sensitisation . In this method, participants indicate themselves whether 20 events happened to them in their childhood. Then, a week later, participants receive five events and are told that they themselves had indicated in the previous week that these events happened to them in their childhood. However, the participatns actually only indicated for four of the events that they were really experienced. The fifth event is an event which participants had indicate *did not* happen to them in the first session. Participants are asked to rate their belief and recollection for all five events, and are given a free recall cue. In line with the mega-analysis of Scoboria and colleagues (2017), studies using the blind implantation paradigm show that around 30% of participants report to have experienced the non-experienced autobiographical event, even though they had previously indicated themselves that this event did not happen to them (Otgaar et al., 2022; Shahvaroughi et al., 2025). The current study examined whether sensitisation instructions can successfully reverse false autobiographical reports (i.e., beliefs or memories) generated in a blind implantation paradigm. This issue will provide important insights into critiques that in classical implantation studies informants might have been wrong and "false memories" reported might actually be true memories (Andrews & Brewin, 2024). Using a blind implantation method without an informant reduces the chances of this happening, and thereby will also provide theoretical insights on whether false memories can be reversed with the source sensitisation method. Moreover, the suggestion in the blind implantation paradigm seemingly stems from the participants themselves, which is a different source than in other false memory paradigms. Source sensitisation training might then help the participant differentiate between memories for events they truly experienced versus ones that came from an internally suggested source. Thus, it is of practical interest whether beneficial effects of source sensitisation extend to other kind of suggestive sources than those tested before (i.e., suggestions from other people during interviews or from written statements; Oeberst et al., 2021; Spearing et al., 2025; Szpitalak et al., 2021). Another issue we addressed in the current study is that sensitisation strategies used in previous research have taught participants about memory sources and/ or false memories in a rather passive way by, for example, providing educational videos (Spearing et al., 2025; Szpitalak et al., 2021) or simply telling participants about the existence of such phenomena (Oeberst et al., 2021). Yet, a more active sensitisation training might be even more effective in reducing false memory reporting. Indeed, in child participants more active memory trainings have proven to be effective in reducing the reporting of false autobiographical information (e.g., Poole & Lindsay, 2002). Thus, we will supplement sensitisation instructions in video form with a short, more active source memory training in the current study.

Last, only Oeberst and colleagues (2021) examined false beliefs in addition to false memories. The other studies looked at reporting of misinformation. To investigate whether our training is is effective in reducing false *memory* or specifically in reducing false *belief* (i.e., the truth status someone attaches to the idea that the event happened to them), we investigated the effects of sensitisation training on the reporting of both separately, as well as on detailed false reports (i.e., false beliefs or false memories for which the participant also provides additional false details in their cued recall).

The current study

Hence, in the current study, we examined how rich false autobiographical memories generated via a blind implantation method might be reduced when source sensitisation strategies are applied. Our false memory reduction strategy was adopted from the strategies used by both Oeberst et al. (2021) and Spearing et al. (2025). Specifically, next to an educational video we also provided participants with three exercise scenarios in which their source monitoring was tested. If participants made a mistake, they were provided with feedback explaining why their answer in the training was incorrect.

Hypotheses

Based on previous literature, we formulated several hypotheses. First, we expected that participants who received the source sensitisation memory training (SMT) would be less likely to provide a (detailed) false report (i.e., 15% compared with 30% in the control group) of an unexperienced autobiographical event compared to participants who did not receive such a source sensitisation training (H1). Second, we expected that those who received the SMT would have fewer false memories of an unexperienced autobiographical event compared with those who did not receive such training (H2). The difference to the first hypothesis lies in the type of false memory (i.e., H1 pertains to *detailed* false reports; H2 pertains to self-rated false memories). Both hypothesis 1 and 2 were based on previous research using sensitisation instructions (e.g., Oeberst et al., 2021; Spearing et al., 2025) and source memory training (e.g., Poole & Lindsay, 2002) as a means to reduce the reporting of false memories. Third, we expected that those receiving the SMT would report lower false beliefs of an unexperienced autobiographical event compared with those who did not receive such training (H3; Oeberst et al., 2021; Poole & Lindsay, 2002; Spearing et al., 2025). Fourth, we did not expect that the SMT would lead to a lower proportion of memories and beliefs for truly experienced events (i.e,. lead to general skepticism).[2] Although direct post warnings have sometimes led to such a "tainted truth" effect in which post warnings did not just reduce reporting of false, but also of true memories, more indirect sensitisation instructions have not had such negative effects on true memory reporting (Oeberst et al., 2021; Spearing et al., 2025; Szpitalak et al., 2021).

Method

Design and participants

The current study used a between subjects design with training 2 (source sensitisation memory training (SMT) versus unrelated training) as the between subjects condition. Participants were randomly assigned to the two groups. For testing of our fourth hypothesis, the design was mixed with the four true events as a within subject factor and the training as a between subjects factor, leading to the following design: 4 (within subjects: true event 1 verus true event 2 versus true event 3 versus true event 4) x 2 (between subjects: SMT versus unrelated training). There were several dependent variables: having a false memory (binary and numerical)[3], having a false belief (binary and numerical), true memory (for an event that did happen), true belief (for an event that did happen), and providing a detailed false report (binary). The variable of interest for the power analysis was whether a detailed false report of the unexperienced autobiographical event was provided by participants.

Effect size considerations

In our hypotheses, we expected a 50% decrease (i.e., from 30% of participants making a false report to 15% making a false report) in false memories/ beliefs/ detailed false reports in the SMT group compared with the control group. This decrease was based on the following rationale: previous research using a sensitisation strategy has found that false memory reports reduced back to baseline (i.e., the same percentage of false memories as reported at the start of the study) after the implementation of two strategies aimed at sensitisation (i.e., false memory sensitisation and source sensitisation – both were employed by educating participants), thereby showing a significant effect of such sensitisation strategies (Oeberst et al., 2021). Other false memory reduction strategies that have been considered "successful" have reduced the number of false memories generated by half (e.g., post-warnings, Blank & Launay, 2014), but cannot be implemented in practice due to constraints of the forensic situation (see above).

If our training can achieve a reduction of detailed false memory reports of 50% (i.e., reducing the reported false memories to 15% from 30%), then this training could be considered effective. The cost of administering this training is somewhat larger than just providing information (i.e., the sensitisation strategy employed by Oeberst et al., 2021 or Spearing et al., 2025) because participants also receive an active training in distinguishing memories from different sources. Therefore, even though also smaller percentages are interesting in terms of whether this SMT helps to reduce rich autobiographical false memories, implementation (based on a cost–benefit consideration) should be recommended if the effect reaches a smallest effect size of interest of a reduction of 15% points (i.e., thereby reducing the reported false memories by half).

Participants

For sample size planning, we ran simulation-based power analyses for logistic regressions (see R code on OSF: https://osf.io/k23hr/?view_only=94814c8662cc4c69b55ddf4e5eb7c177), assuming the proportions of false events reported in each condition as follows: 30% (control group, based on Otgaar et al., 2022), 15% (SMT group). The logistic regression would reach .811 power with $n = 120$ per group ($N_{total} = 240$; sensitivity power analysis).

In order to take this smallest effect size of interest into account with minimum effects testing, we also conducted a power analysis considering that the effect should fully fall outside of the 95% confidence bounds, thus including the latter in the simulation based power analysis (Riesthuis, 2024). Doing so, the logistic regression – based on the same parameters as described above – would reach a statistical power of .75 with $n = 140$ per group ($N_{total} = 280$, sensitivity power analysis). Thus, in order to reliably see whether the effect reaches this threshold of significance, we aimed to test 280 participants.

We recruited 348 participants who signed up for participation via the KU Leuven Experiment Management System (SONA) and were predominantly university students. Of these initial 348 participants, $n = 323$ were eligible to take part in the main parts of the study (sessions 2 and 3).[4] Due to attrition, 294 participants completed session

2, and 280 completed all three sessions. We had to exclude another 4 participants because they failed more than half of the questions about the video in the control condition ($n = 2$), or failed more than half of the source sensitisation training ($n = 2$). Our final sample for the session three data thus was $N = 276$, with $n = 139$ participants randomly assigned to the control condition and $n = 137$ participants randomly assigned to the source sensitisation memory training condition.[5] We report demographics based on the session 2 data set (i.e., thus including participants who dropped out at Session 3 – the reasoning is that false memory frequencies for session 2 were calculated on all participants who completed the session). Participants predominantly identified as women ($n_{women} = 250$, $n_{men} = 42$, $n_{other} = 2$) and were on average 19.78 years old ($SD = 4.4$; range: 18–59 years old). Moreover, most participants were Belgian ($n = 259$). An overview of other nationalities can be found in Table 1.

We received ethical approval from the social and societal ethics committee at KU Leuven (reference number: G-2024-8452). Participation was voluntary, but participants were able to receive compensation in form of course credit (1 SONA credit at the psychology faculty) or by joining a raffle for one of ten 15 euro Amazon vouchers. We preregistered the study on the Open Science Framework (https://osf.io/qhz35) and our hypotheses, materials, data, R code, and an R markdown file can be found there: https://osf.io/k23hr/?view_only=94814c8662cc4c69b55ddf4e5eb7c177.

Procedure

Our procedure is adapted from other blind implantation studies (e.g., Otgaar et al., 2022; Shahvaroughi et al., 2025). The study spanned three short sessions, with a delay between each session (i.e., one week between sessions 1 and 2, and one day between sessions 2 and 3). The study took place fully online via Lime Survey, where sessions two and three were personalised per participant. A visual representation of the procedure can be found in Figure 1.

Session 1

In session 1, participants' baseline memory (i.e., whether participants remember having experienced specific events in their childhood) was measured. To this end, participants received a list of 20 events (see Appendix) and were asked to indicate if they experienced them in their childhood (yes/no). For each answer, they were also asked to indicate their confidence on an 6 point scale ranging from (1)"not very sure" to (6) "very sure".

Table 1. Demographic background.

Nationality	*n*
Belgian	259
Dutch	24
Pakistani	2
Romanian	2
American	1
Irish	1
Lithuanian	1
Turkish	1
Ukrainian	1
Uzbek	1

Session 2

After a one week delay, eligible participants received a link to session 2. This delay was based on previous work using the blind implantation paradigm (Otgaar et al., 2022; Shahvaroughi et al., 2025). Because our expected false report rate was based on prior research, we wanted to mimic the procedures as much as possible. Moreover, false memory implantation is most successful with a week delay between sessions, (Scoboria et al., 2017). The second questionnaire participants received was personalised, including a list of five events from session 1. All participants received instructions telling them that they had previously indicated that **all five events** had happened to them in their childhood. However, this was only true for four of the events. The fifth event was one of the two critical events (i.e., swimming trunks fell off in a swimming pool or breaking something valuable in a store), which participants indicated did not happen to them in session 1. We decided to go with two events to reduce the number of participants who would have to be excluded from participation based on having experienced one of the events as such rates have been high in past research using this paradigm (Otgaar et al., 2022; Shahvaroughi et al., 2025). In a pilot study by Shahvaroughi et al. (2025) participants were asked to rate a number of possible childhood events in terms of plausibility and valence. The two events we picked as our critical events were rated by participants as similar in terms of valence (breaking something: $M = 1.94$, $SD = 1.25$, swimming trunks: $M = 1.84$, $SD = 1.21$) and in terms of plausibility (breaking something: $M = 4.65$, $SD = 2.15$, swimming trunks: $M = 4.71$, $SD = 2.06$; Shahvaroughi et al., 2025). All participants received the false event as the third event. Participants received one of the two critical events, depending on which one of them they had indicated in Session 1 had *not* happened to them. If neither of the critical events had happened to the participant, one of the two was chosen at random. The other four events were randomly chosen for each participant from the pool of events that they had personally indicated in Session 1 happened to them in childhood – the two critical events could not be included in these four. For each event, participants were asked to rate their belief that the event happened to them on an 8 point scale ranging from (1) "definitely did not happen" to (8) "definitely did happen". Subsequently, participants rated their memory for the event on an 8 point scale ranging from (1) "no memory for the event at all" to (8) "clear and complete memory for the event". After completing these belief and recollection ratings, participants engaged in an imagination exercise which was played to

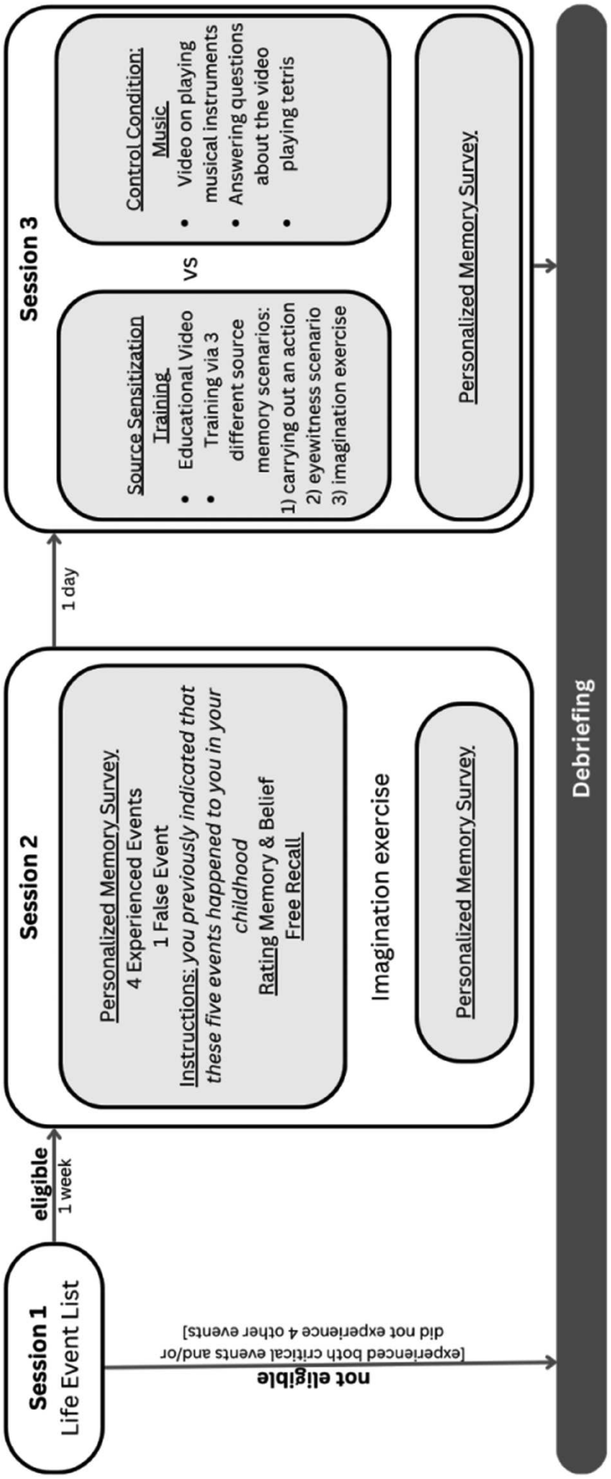

Figure 1. Experimental procedure.

them as an audio message. Specifically, during this imagination exercise participants were told the following:

> Now we want to try a technique that can help you remember events that happened in your childhood. Please sit back, close your eyes and try to imagine [________] in your childhood. Try to imagine yourself at that age, and try to go back to the place where it happened, so try to go back to [________]. Who is with you? When did it happen? What are you feeling? Try to imagine the event for a few moments. When you are done, please open your eyes and click to the next page.

After this imagination exercise, participants were asked to rate their belief and memory one more time.[6] Last, participants were asked to freely recall any details they could remember about the event, giving the following prompts: duration of event, actions taken, people present, age at the time of the event (based on the details included by Loftus & Pickrell, 1995; see also Andrews & Brewin, 2024). Participants completed these steps for all five events.

Session 3

After a one day delay, participants received a third questionnaire, which was personalised to include the same five events from session 2. Previous research using the blind implantation paradigm has not implemented a third session. Yet, other false memory implantation paradigms can consist of several sessions which typically occur in the course of one week (Scoboria et al., 2017). We did not expect that a longer delay would be needed between sessions 2 and 3, as the intent of the third session was not to further strengthen false memories or beliefs, but to examine the possibility of reversing them. Therefore, we decided on a one-day delay between sessions 2 and 3. Participants in the SMT condition received a sensitisation source memory training (see materials on OSF; https://osf.io/k23hr/?view_only=94814c8662cc4c69b55ddf4e5eb7c177). During this training, participants were provided with information about different sources of memories and the existence of false memories (i.e., a combination of a source and false memory sensitisation). Then, they engaged in three small exercises to train identification of memory sources: in the first training task, participants watched a video clip of a person carrying out an action (i.e., cleaning out the coffee capsules from a coffee machine) and stating other actions they sometimes do (i.e., cleaning the machine by pressing a button). Then, participants had to answer questions about the different actions (e.g., "Did you see the person in the video clean out the coffee capsules?" "What is the source of this information?"). This first exercise aimed at differentiating witnessed and heard about actions. In the second training task, participants watched a short video of a theft (Greene et al., 2022), and received an eyewitness account of the theft. Then, participants were instructed to answer questions about the theft, and identify details that were in the video versus only reported by the witness. The second exercise thus aimed at differentiating witnessed events from those others discussed and identifying misleading post event information. In the last task, participants were asked to view two short videos of a family taking a walk in nature, before engaging in an imagery exercise in which they are asked to imagine a dog being present during the walk. Later, participants had to answer some questions about details regarding the videos, and identify if they had seen them or imagined them (e.g., "Did you see a dog in the video?" "What is the source of this information?"). Participants in the control group received a training on what playing musical instruments does to the brain and played Tetris for five minutes. All together, the trainings were comparable in terms of timing.

After the training, all participants were asked about the five events from session 2 one more time. Specifically, for each event participants were asked to provide belief and recollection ratings again (on the same scale as in session 2). After having provided these for all five events, participants were informed that one of the five events was false, and they were asked to guess which event was the false one. They also rated their confidence that the event they mentioned was the false event on a scale from (1) "not confident at all" to (8) "very confident". As the last step, all participants receive a debriefing.

Data analysis

Scoring

Our scoring was based on Otgaar and colleagues' (2022) and Shahvaroughi and colleagues' (2025) scoring, which is the standard used in blind implantation studies and therefore allows for comparison across studies. A false belief was scored when participants indicated a belief score of 5 or higher for the critical event (i.e., breaking something valuable in a store or your swimming trunks fell off at a swimming pool) and recollection score that is lower than the belief score. A false memory was scored when participants have a recollection of 5 or higher and a belief score that is equal to the recollection score. A detailed false report was scored when a false belief or memory is accompanied with any additional details (i.e., beyond the details suggested in the event title). We deemed any additional detail (e.g., location, persons involved, emotion at the time, etc.) sufficient to conclude a detailed false report. Important to note, participants were not required to include all details from the original event description because some of this information was provided in the question. The same strategy was used for true beliefs and memories. We scored true beliefs, true memories, and detailed true reports in the same way.

Software

All data analyses were conducted using the open-source software R (R Core Team, 2023). The following software packages were used: *readxl* (Wickham & Bryan, 2019), *ggplot2* (Wickham, 2016), *psych* (Revelle, 2018), *dplyr* (Wickham et al., 2018), doBy (Halekoh & Højsgaard, 2024), *ggpubr* (Kassambara, 2023), *sjstats* (Lüdecke, 2024), gt (Iannone et al., 2024), *tidyr* (Wickham et al., 2019), *knitr* (Xie, 2018; 2024), *broom* (Robinson et al., 2024), *broom.-mixed* (Bolker & Robinson, 2024), *exact2 × 2* (Fay, 2010), *effsize* (Torchiano, 2020), and *lme4* (Bates et al., 2015). Anonymous data and analysis code (i.e., R Markdown file) can be found on the Open Science Framework: (https://osf.io/k23hr/?view_only=94814c8662cc4c69b55ddf4e5eb7c177). We used OpenAI (2023) to help resolve error messages in code and to annotate R code with comments, to facilitate other researchers' understanding of the steps of the code and thereby increases replicability and transparency.

Results

We used an alpha of 5% for our analyses, because we considered a false positive rate of 5% acceptable for this study. For session 2, we report memory and belief scores obtained *after* the imagination exercise as this exercise was used to strengthen false memories and all participants received it.

False memory

Session 2 (pre-registered)

In the second session, 18.4% (54/ 294) of participants made a false report of the unexperienced event – stemming from 15% (44/294) of participants who reported a false belief and 3.4% (10/294) who reported a false memory. Moreover, 9.2% (27/294) of participants made a *detailed false report,* meaning that next to having a false memory or belief they also came up with extra details about the event in the free recall exercise.

Session 3 (pre-registered)

For the third session, participants were divided into the source sensitisation memory training (SMT) condition and the control condition as participants received either the training or the control training at the start of the third session. 25.5% (35/137) of participants in the SMT condition made a false report of the unexperienced event after the memory training – 20.4% (28/137) reported a false belief and 5.1% (7/137) reported a false memory. Moreover, 13.1% (18/137) made a detailed false report, providing extra details regarding the false event in the free recall exercise on top of self-reporting a false belief or memory. In the control condition, around 25.2% (35/139) of participants made a false report of the unexperienced event. Specifically, a false belief was reported by 22.3% (31/139) of participants in this group, and a false memory by 2.9% (4/139). A detailed false report was given by 11.5% (16/139) of participants in the control condition. Figure 2 is a visual representation of the false memory data from sessions 2 and 3.

Hypothesis 1 (pre-registered)

We ran a logistic regression to examine whether the source sensitisation memory training (versus not receiving the training) would impact the proportion of participants who reported a detailed false account of the non-experienced event at the end of the third session. The training

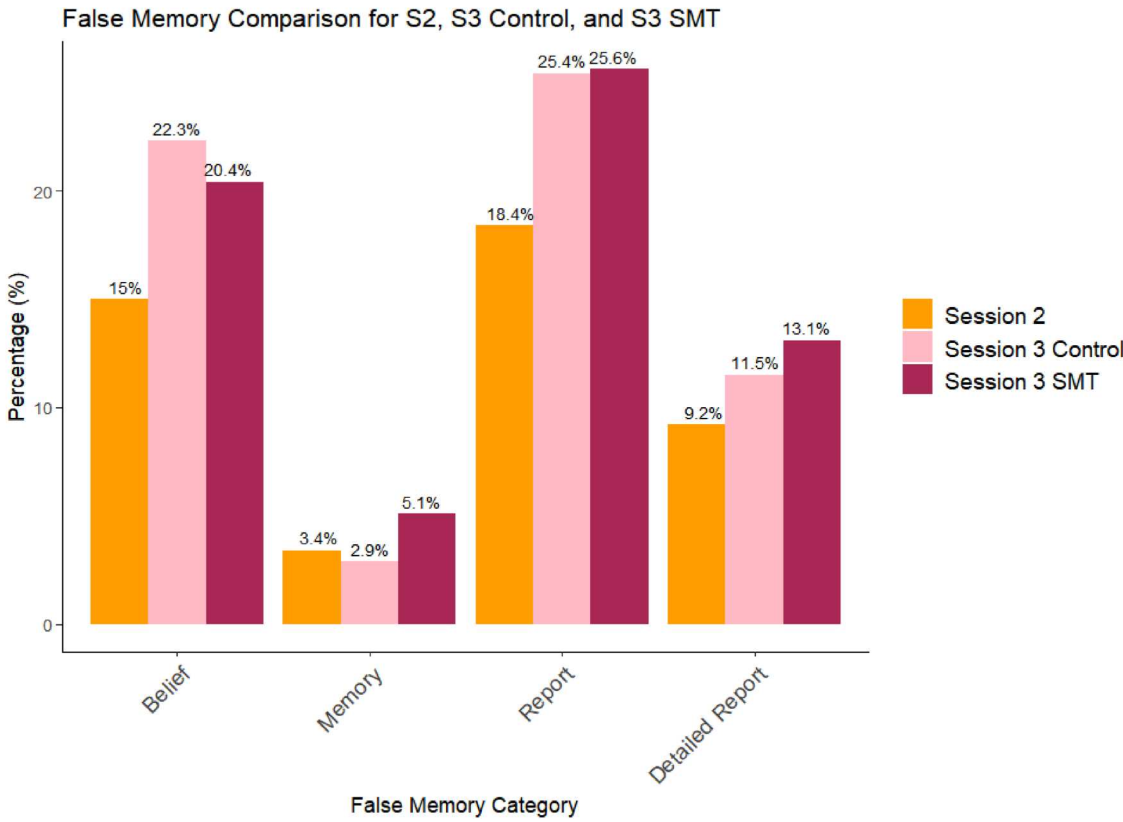

Figure 2. False memory comparison sessions 2 and 3 [SMT vs Control].

was not a statistically significant predictor of whether or not participants made a detailed false report, $\beta = .15$, 95% CI [−.57; .88], $p = .68$. Because we set a SESOI (of a difference of .15), we also completed an equivalence test to examine whether this effect is not just not statistically insignificant but also equivalent. To this end, we converted our probability based SESOI interval [−.15; .15] into log odds to match the effect size of our model. This leaves us with SESOI bounds in log odds of [−.89; .89]. If the achieved 90% CI of the effect of source sensitisation training on the probability of making a detailed false report falls fully within these SESOI bounds the effect is equivalent (Riesthuis, 2024). This was the case, as we found $\beta = .15$, 90% CI [−.45; .76]. Thus, the effect was equivalent.

Hypothesis 2 (pre-registered)

We also conducted a logistic regression to examine whether our training had any impact on whether participants reported a false memory for the unexperienced event or not. The training was not a statistically significant predictor of whether participants reported a false memory, $\beta = .56$, 95% CI [−.62; 1.96], $p = .35$. Here we again also performed an equivalence test using the same 15% SESOI and log odd bounds [−.89; .89]. The 90% CI of our effect [−.43; 1.72] overlapped with the upper bound, meaning that the result is inconclusive.

Hypothesis 3 (pre-registered)

Next, we ran a logistic regression to examine whether our training had an impact on whether participants reported a false belief of the unexperienced event. Again, the training was not a statistically significant predictor of participants reporting a false belief, $\beta = -.11$, 95% CI [−.69; .47], $p = .71$. We performed an equivalence test using our 15% SESOI log odd bounds [−.89; .89], and the 90% CI of our predictor (i.e., receiving the source sensitisation training) fell completely into the bounds of the SESOI, 90% CI [−.60; .37]. This meant that the effect was statistically equivalent.

Picking the false event (not pre-registered)

At the end of the third session, we debriefed participants and asked them to indicate which of the five events they thought was the false one. In the control condition, 61.9% (86/139) correctly identified the false event, while the other 38.1% picked an event they had initially indicated happened to them (in S1) as the false event. In the SMT group 69.3% (95/137) correctly identified the false event, whereas 30.7% (42/137) identified one of the true events as a false one. We ran a logistic regression with training condition as the predictor (SMT versus control) and correctly identifying the false event at the end of session 3 as the dependent variable and found no statistically significant difference in the proportion of

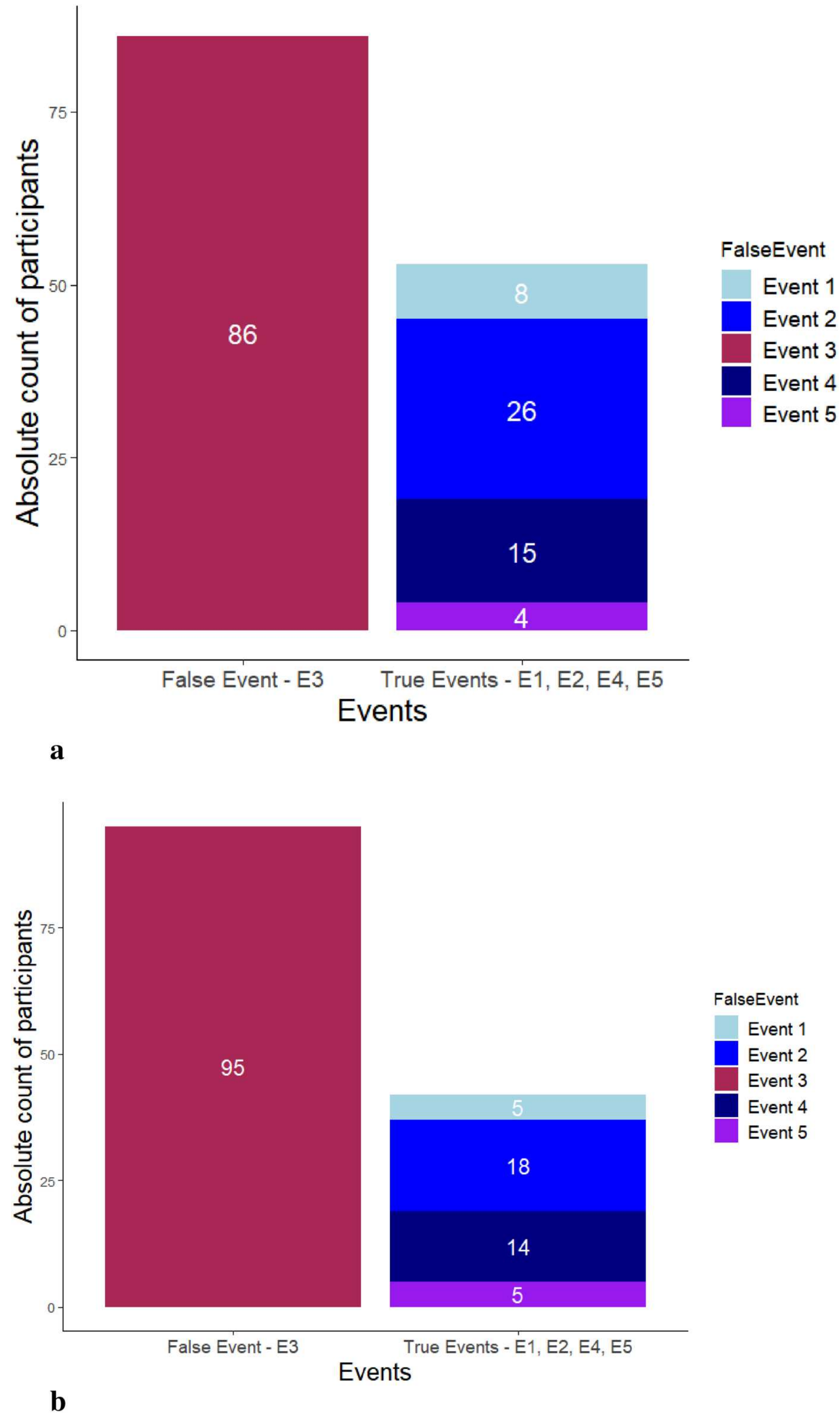

Figure 3. a. Events picked by participants in the SMT condition as the false event. b. Events picked by participants in the SMT condition as the false event.

participants who correctly identified the false event after debriefing across the two training conditions, $\beta = .33$, 95% CI [−.17, .83], $p = .19$. Figure 3a and b visually represent the spread of events picked by participants in the two conditions as the false event.

We further asked participants about how confident they were in their choice and conducted a Welch independent samples t-test to examine whether participants who had correctly identified the false event (confidence: $M = 5.46$, $SD = 1.98$) compared to those who picked a true event as the false one (confidence: $M = 5.23$, $SD = 2.33$) differed in how confident they were in their choice. We found no evidence of a statistically significant difference, $t(166.21) = .81$, $p = .42$, $d = −.11$, 95%CI [−.36; .14].

False memory across sessions (not pre-registered)

We were interested in whether participants had (statistically significantly) increased their self-reported false belief and false memory scores (numerical) from session 2 to session 3. To this end, we performed two paired samples t-tests. Participants statistically significantly increased their false belief in the non-experienced event from session 2 ($M = 2.71$; $SD = 2.1$) to session 3 ($M = 3.0$; $SD = 2.2$), $t(278) = −3.38$, $p < .001$, *mean difference* = −.29, 95% CI [−.46; −.12], $d = .14$, 95% CI [.06; .21]. However, false memory scores did not change statistically significantly from session 2 ($M = 1.76$; $SD = 1.72$) to session 3 ($M = 1.82$; $SD = 1.73$), $t(278) = −.87$, $p = .39$, *mean difference* = −.06 9, 95% CI [−.2; .08], $d = −.04$, 95% CI [−.12; .05]. We further also analysed whether participants' likelihood to provide a detailed false report (categorical) increased from session 2 to session 3 by conducting a McNemar test. McNemar's test takes into account the paired nature of our data (i.e., each participant has a score for session 2 and for session 3). We found evidence that the likelihood to provide a detailed false report increased statistically significantly from Session 2 to Session 3, X^2 (279, 1) = 4.05, $p = .044$, $\phi = 0.45$. Of note, we only included participants in this analysis who completed both sessions 2 and 3.

Confidence at session 1 (not pre-registered)

We wanted to investigate whether how confident participants were that the critical event did *not* happen to them in the first session affected whether they provided a detailed false report in the last session (session 3). To this end, we performed a logistic regression examining the effect of confidence score for the critical event in session 1 on the likelihood to provide a detailed false report in session 3. We found that *lower* confidence in session 1 that the critical event was *not* experienced was statistically significantly related to a *higher* likelihood to make a detailed false report in session 3, $\beta = −.29$, 95% CI [−.53; −.05], $p = .016$.

Table 2. Remembering true events session 2.

Event		Belief		Memory		Detailed Report	
		n	%	*n*	%	*n*	%
Session 2							
True Event 1		188	63.9	93	31.6	271	92.2
True Event 2		171	58.2	96	32.7	250	85.0
True Event 3		156	53.1	116	39.5	249	84.7
True Event 4		147	50.0	137	46.6	265	90.1
Session 3							
True Event 1	*SMT*	90	65.7	42	30.7	128	93.4
	Control	80	57.6	49	35.5	121	87.1
True Event 2	*SMT*	85	62.0	44	32.1	116	84.7
	Control	81	58.3	43	30.9	111	79.9
True Event 3	*SMT*	70	51.1	58	42.3	114	83.2
	Control	81	58.3	46	33.1	116	83.5
True Event 4	*SMT*	75	54.7	58	42.3	122	89.1
	Control	67	48.2	67	48.2	121	87.1

Note: Total participant number for session 2 is $N = 294$, for session 3 it is $n = 137$ in the SMT condition and $n = 139$ in the control condition (i.e., total 276 participants in session 3).

True memory

Session 2 (pre-registered)

We also examined how well participants remembered the four true events. 95.6% (281/294) of participants reported a belief (63.9%; 188/294) or a memory (31.6%; 93/294) for the first true event, and 92.2% (271/294) reported additional details about the first true event. Because percentages for the four true events are relatively similar, we do not report all percentages in text – however, percentages for all four true events are reported in Table 2.

Session 3 (pre-registered)

96.4% (132/137) of participants in the SMT condition made a report (i.e., reported having either a memory or belief) of the first true event in session 3–65.7% (90/137) reported a belief and 30.7% (42/137) reported a memory. Additionally, 93.4% (128/137) provided extra details in the free recall exercise. In the control condition, around 92.9% (129139) of participants made a report of the first true event in S3. Specifically, a belief was reported by 57.6% (80/139) of participants in the control group, a memory by 35.3% (49/139). A detailed report was given by 87.1% (121/139) of participants in this group. Memory scores for the other true events are reported in Table 2 (see rationale above).

Hypothesis 4 (pre-registered)

Last, we examined whether our training led to a "tainted truth effect" – a decrease in true memory reporting after participants took part in the training compared to not taking part in the training. Because we considered all four true events, we ran a multilevel logistic regression taking into account that each participant provided ratings on four true events. Thus, the participant number posed as the random effect, whereas the training condition served as our fixed predictor of participants' true memory score. Condition was not a statistically significant predictor of whether participants reported true memories, $\beta = −0.04$, 95% CI [−.61; .54], 90% CI [−.52; .45], $p = .90$. Moreover, using the 15% SESOI bounds of [−.89; .89] log odds, the effect was equivalent, as the 90% CI [−.52; .45] fit fully into these SESOI bounds.

We also examined the tainted truth effect by running a similar multilevel logistic regression analysis on the outcome variable of "true belief", with participant as a random effect and source memory training as the fixed effect. Source memory training was not a statistically significant predictor of having a true belief, $\beta = .20$, 95% CI [−.34; .74], 90% CI [−.26; .65], $p = .48$. The effect was also equivalent, as our 90% CI [−.26; .65] fit fully within the bounds of our SESOI [−.89; .89].

Discussion

The main aim of the current study was to examine the effectiveness of a source sensitisation memory training in reducing or reversing false autobiographical memories. We found that source sensitisation memory training had no effect on the reduction of false memories, nor on participants' ability to correctly identify which event was the false one after debriefing. Moreover, irrespective of receiving sensitisation training, the likelihood of making a detailed false report increased over time, from session 2 to session 3. We discuss these and other findings below.

Despite our expectations that the source sensitisation memory training would reduce the creation of false beliefs, false memories, and detailed false reports, we found no effect of the training on the likelihood of making such false reports. Previous research using similar sensitisation techniques (albeit without training) had found some success in reducing the report of false details suggested by another person (Spearing et al., 2025) and the likelihood of forming rich false memories for non-experienced events (Oeberst et al., 2021). Furthermore, a memory training aimed at increasing source monitoring skills has been successful with reducing the reporting of misinformation in children (Poole & Lindsay, 2002). Surprisingly, our sensitisation instructions paired with a source monitoring skill training did not yield similar, or even better results. There are several possible explanations for our finding.

First, it is possible that false reports elicited with the blind implantation paradigm are difficult to identify for participants because the suggestion seemingly does not come from an external source. Indeed, in a blind implantation procedure, the suggestion that the unexperienced event actually happened to the participant appears to come from the participant themselves. Specifically, we told participants that they themselves had indicated in the first session that the event happened to them. One possibility is that the self-generated suggestion made the event more plausible and self-relevant to the participant, which potentially increased the likelihood to make a false report (Scoboria et al., 2017; 2004) and might ostensibly also have prevented reduction of a false report or belief. More research on false memories in the blind implantation paradigm is needed to examine the robustness of this kind of false report. For example, such research could directly compare false belief and false memory rates when it is suggested that the participants themselves indicated initially that the event happened to them versus telling them that their parents had indicated so.

Another possible explanation is that our source sensitisation memory training was too easy. That is, our training included basic source monitoring skills that might arguably already be developed in adults. Previous research using a source memory training to reduce false reporting was conducted on children who do not yet have fully developed source monitoring skills (Poole & Lindsay, 2002). If the training we employed was not challenging enough it might not have been clear to participants what the aim of the training was, rendering it difficult to transfer any newly learned skills to the memory test at the end of session 3 (Bjork & Bjork, 2020). In contrast to previous research which found that source sensitisation instruction reduced false memory reporting (Oeberst et al., 2021; Spearing et al., 2025; Szpitalak et al., 2021), we actually found inconclusive evidence in our equivalence test, and it seemed like our training possibly increased false memory reports – thus having the opposite effect than intended. If the training was too easy, it could have made participants overconfident (Schraw & Roedel, 1994) that they would notice had they been presented with false information – as source monitoring then might have seemed like an easy skill to them. This, in turn, could have blocked potential positive effects of the sensitisation instructions participants were presented with prior to the training section. Future research could aim to disentangle this by presenting some participants only with the sensitisation instructions, some with the instructions and training, and some with a control condition to examine if the beneficial effects of sensitisation instructions found in previous research can be replicated when the suggestion seems to stem from the person themselves (i.e., as is the case in the blind implantation paradigm). Future research employing related trainings could also include a rating of participants of how difficult they found the training versus the final steps of the experiment in order to clarify whether the training was received appropriately.

Another possibility is that the training – although each step by itself might have been easy – included too many steps rendering its application and transfer difficult. Participants might have tried to focus on too many elements of the training simultaneously during the last belief and recollection ratings, or failed to understand how to concretely apply the skills in the training to the memory exercise. Indeed, other studies that employed educational tools in combination with a several-step training programme have found that covering too many aspects in the training can be overwhelming and reduce training effectiveness – even when each training step by itself was well received by participants. For instance, Seabrooke and colleagues (2025) found that participants were able to apply learned skills better in a following task if they underwent a shorter training focused only on one skill rather than a longer more encompassing training. Another example is that Oeberst and colleagues (2021) tested memory after each step of their sensitisation instructions: once after providing source sensitisation instructions (i.e., that our memories can stem from different sources), and then again after providing false memory sensitisation instructions (i.e., specifically that we can have false memories and where they stem from), thereby allowing participants to employ each newly learned information separately. Future studies could investigate if a shortened version of our training is more successful in reducing false reports of autobiographical memories, or if a step-by-step method is successful in reducing false reports.

Even though our training did not have the predicted positive effects, it also did not have any negative effects on memory: the training did not affect reporting of true memories, in line with our expectations (H4). A *tainted truth effect* (Echterhoff et al., 2007) has especially been found in the past when direct warnings were used, whereas with more indirect warnings or instructions such an effect has usually been absent (Spearing et al., 2025; Szpitalak et al., 2021). Thus, our participants did not develop a general skepticism towards their own memories reported throughout the experiment. Given that participants did not become more skeptical of the false event (i.e., the likelihood of reporting the false event was not affected by the training), it would have been very surprising if they had become more skeptical of their true memories based on this same training.

Interestingly, we found that irrespective of training, participants' likelihood to make a detailed false report was larger in session 3 than in session 2. Moreover, participants on average increased their belief that the false event happened to them from session 2 to session 3 – although this increase was only .3 points on the eight point belief scale. Relatedly, we found a Cohen's *d* of .14, meaning that participants on average had a belief score in session 3 that was 0.13 standard deviations higher than in session 2 (Riesthuis et al., 2025). In terms of probability of superiority this means that there is a .54 probability that a randomly chosen participant at session 3 has a higher score than a randomly chosen participants at session 2. While the belief did not change by much, participants were more likely to provide extra details during the free recall in session 3 compared with session 2. It is generally not uncommon for the likelihood to report a false autobiographical memory to increase when exposed to more sessions and thus suggestions about the false event (Scoboria et al., 2017). Similarly, in Oeberst and colleagues (2021), three (suggestive) interviews were conducted with participants in order to implant the false event in their memory, and the percentages of participants who made such false reports increased with each interview until sensitisation instructions were introduced in interviews four and five. The current study showed for the first time that the same occurs in the blind implantation paradigm when several sessions are employed. More research should look into repeated suggestions in the blind implantation paradigm to assess the upper limit of false memory implantation in this paradigm.

Another important finding of the current study was that participants who were more confident during session 1 that the critical event had *not* happened to them were less likely to make a detailed false report later on (i.e., in session 3) compared to participants who were initially less confident that they had never experienced the critical event(s). This suggests that uncertainty allows room for participants to form false autobiographical memories. Perhaps participants who were less sure about whether they had (not) experienced the critical event saw the event as something that plausibly could have happened to them, thereby increasing chances of accepting that the suggested event happened to them (Scoboria et al., 2017; 2004) and making a detailed false report later on. Of course, it is also possible that participants who were less confident that they had *not* experienced the event were people who actually did experience the event but did not remember this or made a mistake at session 1 when indicating which events had or had not happened to them. The idea that a subset of participants who report false memories in implantation studies actually did experience the event and are reporting true memories is a common criticism of implantation studies (Andrews & Brewin, 2024; see also: Murphy & Greene, 2025). Although in classic implantation studies this concern stems from parents' being informants for whether participants experienced certain events, the same limitation might also apply in the blind implantation paradigm when participants themselves have forgotten that an event happened to them, and only remember it in later sessions, seemingly leading to a detailed false report.

Conclusion

With the current study we aimed to test the effectiveness of a source sensitisation training in reducing false autobiographical memories. However, our training was not effective in reducing reports of false autobiographical memories, but it also did not negatively affect remembering of true events. The current study also yielded some insights on false reports elicited via blind implantation, as we found that repeated suggestions that the participant had indicated the event to have happened to them increased false reports, and that uncertainty surrounding whether the event happened fostered later false memory reports. Further research is needed to test the effectiveness of a shortened version of the training in reducing the likelihood to report a false autobiographical memory, and whether the beneficial effects of sensitisation instructions generalise to false reports generated in the blind implantation paradigm.

Notes

1. Oeberst et al. (2021) scored memory and belief on one continuous scale of 1 – rejecting the event, 2 – no memory, 3 – accepting event (i.e., false belief), 4 – partial (false) memory, 5 – full (false) memory, 6 – robust (false) memory
2. Here we diverged slightly from the preregistration: in the preregistration we formulated the hypothesis as follows: "Fourth, we do not expect to find a tainted truth effect (Szpitalak et al., 2021; Oeberst et al., 2021), meaning that we do not expect that the SMT leads to lower memory and belief scores (numerical) for truly experienced events (i.e., leads to general skepticism)". However, while conducting the analysis we decided it would make more sense to score true memories and beliefs the same way in which we scored false memories/ beliefs and

thus as categorical variables. The sentiment of the hypothesis remains the same as preregistered.
3. See scoring for additional information on how these variables are coded.
4. Participants had to fulfil the following criteria regarding eligibility. First, participants either had to not have experienced critical event 1 (i.e., breaking something valuable in a store) or critical event 2 (i.e., your swimming trunks fell off at a swimming pool). Moreover, participants had to have experienced at least 4 events from the list of events other than critical event 1 and 2. Twenty-five participants had to be excluded due to these criteria.
5. Of note, the two critical items were used with similar frequency, with $n = 166$ participants receiving the prompt for loosing swimming trunks while swimming in a public pool, and $n = 153$ participants receiving the prompt that they had broken something valuable in a store [both pertaining to childhood memories].
6. These were the memory and belief ratings we used to assess true and false memory scores in session 2.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The current paper was supported by a Fonds Wetenschappelijk Onderzoek research project awarded to the last author [grant number: G0D3621N] and a Fonds Wetenschappelijk Onderzoek post-doctoral fellowship awarded to the second author [grant number: 1203824N].

Data availability statement

All materials, anonymized data, R code, and R markdown files can be found on the Open Science Framework at: https://osf.io/k23hr/?view_only=94814c8662cc4c69b55ddf4e5eb7c177 or under the following DOI: 10.17605/OSF.IO/K23HR.

ORCID

Charlotte A. Bücken https://orcid.org/0000-0001-5128-8348
Paul Riesthuis http://orcid.org/0000-0001-6520-2453

References

Andrews, B., & Brewin, C. R. (2024). Lost in the mall? Interrogating judgements of false memory. *Applied Cognitive Psychology*, *38*(6), e70012. https://doi.org/10.1002/acp.70012
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01
Bernstein, D. M., & Loftus, E. F. (2009). The consequences of false memories for food preferences and choices. *Perspectives on Psychological Science*, *4*(2), 135–139. https://doi.org/10.1111/j.1745-6924.2009.01113.x
Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, *9*(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003
Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, *3*(2), 77–88. https://doi.org/10.1037/h0101798
Bolker, B., & Robinson, D. (2024). _broom.mixed: Tidying methods for mixed models_. R package version 0.2.9.6. https://CRAN.R-project.org/package=broom.mixed
Echterhoff, G., Groll, S., & Hirst, W. (2007). Tainted truth: Overcorrection for misinformation influence on eyewitness memory. *Social Cognition*, *25*(3), 367–409. https://doi.org/10.1521/soco.2007.25.3.367
Fay, M. P. (2010). Two-sided exact tests and matching confidence intervals for discrete data. *R Journal*, *2*(1), 53–58. https://digitalcommons.unl.edu/r-journal/313/
Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, *3*(2), 208–214. https://doi.org/10.3758/BF03212420
Geraerts, E., Bernstein, D. M., Merckelbach, H., Linders, C., Raymaekers, L., & Loftus, E. F. (2008). Lasting false beliefs and their behavioral consequences. Psychological Science, 19(8), 749–753. https://doi.org/10.1111/j.1467-9280.2008.02151.x
Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, *26*(1), 20–33. https://doi.org/10.3758/bf03211367
Goodman, G. S., Wang, Y., & Mitchell, C. (2025). The "false memory" implantation paradigm: Commentary on andrews and brewin. *Applied Cognitive Psychology*, *39*(4), e70091. https://doi.org/10.1002/acp.70091
Greene, C. M., Bradshaw, R., Huston, C., & Murphy, G. (2022). The medium and the message: Comparing the effectiveness of six methods of misinformation delivery in an eyewitness memory paradigm. *Journal of Experimental Psychology: Applied*, *28*(4), 677–693. https://doi.org/10.1037/xap0000364
Greene, C. M., Ryan, K. M., Ballantyne, L., Barrett, E., Cowman, C. S., Dawson, C. A., Huston, C., Maher, J., & Murphy, G. (2024). Unringing the bell: Successful debriefing following a rich false memory study. *Memory & Cognition*, *52*(5), 1079–1092. https://doi.org/10.3758/s13421-024-01524-9
Halekoh, U., & Højsgaard, S. (2024). doBy: Groupwise statistics, LSmeans, linear estimates, utilities. R package version 4.6.24. https://CRAN.R-project.org/package=doBy
Herman, S. (2010). The role of corroborative evidence in child sexual abuse evaluations. *Journal of Investigative Psychology and Offender Profiling*, *7*(3), 189–212. https://doi.org/10.1002/jip.122
Iannone, R., Cheng, J., Schloerke, B., Hughes, E., Lauer, A., Seo, J., Brevoort, K., & Roy, O. (2024). _gt: Easily create presentation-ready display tables_. R package version 0.11.1. https://CRAN.R-project.org/package=gt
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3
Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85. https://doi.org/10.1037/0033-295x.88.1.67
Kassambara, A. (2023). ggpubr: 'ggplot2' based publication ready plots. R package version 0.6.0. https://CRAN.R-project.org/package=ggpubr
Kloft, L., Monds, L. A., Blokland, A., Ramaekers, J. G., & Otgaar, H. (2021). Hazy memories in the courtroom: A review of alcohol and other drug effects on false memory and suggestibility. *Neuroscience & Biobehavioral Reviews*, *124*, 291–307. https://doi.org/10.1016/j.neubiorev.2021.02.012
Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2011). *Tell me what happened: Structured investigative interviews of child victims and witnesses* (Vol. 56). John Wiley & Sons.
Leander, L. (2010). Police interviews with child sexual abuse victims: Patterns of reporting, avoidance and denial. *Child Abuse & Neglect*, *34*(3), 192–205. https://doi.org/10.1016/j.chiabu.2009.09.011
Loftus, E. F., & Klemfuss, J. Z. (2023). Misinformation – past, present, and future. *Psychology, Crime & Law*, *30*(4), 312–318. https://doi.org/10.1080/1068316X.2023.2219813

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false mfemories. *Psychiatric Annals*, *25*(12), 720–725. https://doi.org/10.3928/0048-5713-19951201-07
Lüdecke, D. (2024). sjstats: Statistical functions for regression models (Version 0.19.0). https://doi.org/10.5281/zenodo.1284472. https://CRAN.R-project.org/package=sjstats
Lyon, T. D., Williams, S., & Stolzenberg, S. N. (2020). Understanding expert testimony on child sexual abuse denial after New Jersey v. J.L.G.: Ground truth, disclosure suspicion bias, and disclosure substantiation bias. *Behavioral Sciences & the Law*, *38*(6), 630–647. https://doi.org/10.1002/bsl.2490
Murphy, G., Dawson, C. A., Huston, C., Ballantyne, L., Barrett, E., Cowman, C. S., Fitzsimons, C., Maher, J., Ryan, K. M., & Greene, C. M. (2023). Lost in the mall again: A preregistered replication and extension of loftus & pickrell (1995). *Memory (Hove, England)*, *31*(6), 818–830. https://doi.org/10.1080/09658211.2023.2198327
Murphy, G., & Greene, C. M. (2025). Who got lost in the mall? Challenges in counting and classifying false memories. *Applied Cognitive Psychology*, *39*(2), e70044. https://doi.org/10.1002/acp.70044
Nichols, R. M., & Loftus, E. F. (2019). Who is susceptible in three false memory tasks? *Memory (Hove, England)*, *27*(7), 962–984. https://doi.org/10.1080/09658211.2019.1611862
Oeberst, A., Wachendörfer, M. M., Imhoff, R., & Blank, H. (2021). Rich false memories of autobiographical events can be reversed. *Proceedings of the National Academy of Sciences*, *118*(13), e2026447118. https://doi.org/10.1073/pnas.2026447118
OpenAI. (2023). ChatGPT (May 20th 2025 version) [Large language model]. https://chat.openai.com/chat
Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect*, *24*(6), 733–752. https://doi.org/10.1016/s0145-2134(00)00137-x
Otgaar, H., Moldoveanu, G., Melis, V., & Howe, M. L. (2022). A new method to implant false autobiographical memories: Blind implantation. *Journal of Applied Research in Memory and Cognition*, *11*(4), 580. https://doi.org/10.1037/mac0000028
Poole, D. A., & Lindsay, D. S. (2002). Reducing child witnesses' false reports of misinformation from parents. *Journal of Experimental Child Psychology*, *81*(2), 117–140. https://doi.org/10.1006/jecp.2001.2648
R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
Revelle, W. (2018). psych: Procedures for psychological, psychometric, and personality research. https://CRAN.R-project.org/package=psych
Riesthuis, P. (2024). Simulation-based power analyses for the smallest effect size of interest: A confidence-interval approach for minimum-effect and equivalence testing. *Advances in Methods and Practices in Psychological Science*, *7*(2), 1–14. https://doi.org/10.1177/25152459241240722
Riesthuis, P., Howe, M. L., & Otgaar, H. (2025). Meaningful approaches to assessing the size of effects in memory research: Applications and recommendations for study design, interpretation, and analysis. *Memory (Hove, England)*, *33*(5), 485–494. https://doi.org/10.1080/09658211.2025.2492496
Robinson, D., Hayes, A., & Couch, S. (2024). _broom: Convert statistical objects into tidy tibbles_. R package version 1.0.7. https://CRAN.R-project.org/package=broom
Schraw, G., & Roedel, T. D. (1994). Test difficulty and judgment bias. *Memory & Cognition*, *22*(1), 63–69. https://doi.org/10.3758/BF03202762
Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology*, *18*(7), 791–807. https://doi.org/10.1002/acp.1062
Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., & Hyman, I. E. (2017). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory (Hove, England)*, *25*(2), 146–163. https://doi.org/10.1080/09658211.2016.1260747
Seabrooke, T., Pattni, M., & Higham, P. A. (2025). Enhancing human detection of real and AI-generated hyperrealistic faces. *PsyArXiv Preprints on OSF*. https://doi.org/10.31234/osf.io/xjemh_v2.
Shahvaroughi, A., Dyevre, A., & Otgaar, H. (2025). Examining the effects of mood and emotional valence on the creation of false autobiographical memories. *Memory & Cognition*, 1–15. https://doi.org/10.3758/s13421-025-01697-x
Spearing, E. R., Mah, E. Y., Jagota, R., Wade, K. A., Blank, H., & Lindsay, D. S. (2025). Sensitization instructions can reduce the misinformation effect and improve the eyewitness confidence – accuracy relationship. *Journal of Applied Research in Memory and Cognition*, *14*(2), 178–193. https://doi.org/10.1037/mac0000186
Szpitalak, M., Woltmann, A., Polczyk, R., & Kękuś, M. (2021). Memory training as a method for reducing the misinformation effect. *Current Psychology*, *40*(11), 5410–5419. https://doi.org/10.1007/s12144-019-00490-9
Torchiano, M. (2020). effsize: Efficient effect size computation.
Vredeveldt, A., van Rosmalen, E. A. J., van Koppen, P. J., Dror, I. E., & Otgaar, H. (2022). Legal psychologists as experts: Guidelines for minimizing bias. *Psychology, Crime & Law*, *30*(7), 705–729. https://doi.org/10.1080/1068316X.2022.2114476
Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer Cham.
Wickham, H., Averick, M., Bryan, J., Chang, W., Mcgowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. The Journal of Open Source Software, 4(43), 1686. https://doi.org/10.21105/joss.01686
Wickham, H., & Bryan, J. (2019). Readxl: Read excel files. *R package version*, *1*(1), 785. https://CRAN.R-project.org/package=readxl
Wickham, H., François, R., Henry, L., & Müller, K. (2018). dplyr: A grammar of data manipulation. R package version 0.7. https://CRAN.R-project.org/package=dplyr
Xie, Y. (2018). Knitr: A comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible research* (pp. 3–31). Chapman and Hall/CRC.
Xie, Y. (2024). _knitr: A General-purpose package for dynamic report generation in R_. R package version 1.49. https://yihui.org/knitr/

Appendix

List of autobiographical events:

1) Going to a theme park.
2) Experiencing a car crash.
3) swimming trunks fell off at the pool*
4) Undergoing surgery while hospitalised.
5) Having a tooth extracted by a dentist.
6) Building camps with friends.
7) Going ice skating on natural ice in the winter.
8) Going camping with your family in tents.
9) Wounding yourself by falling off your bike.
10) Having a verbal fight with your friends at school.
11) Going on a hot air balloon ride.
12) Meeting your favourite figure at Disneyland Paris.
13) Breaking something valuable in a store*
14) Going to the doctor to suture a wound.
15) Breaking something while falling of a trampoline.
16) A plaster of a broken arm or leg written full by classmates, family, etc.
17) Going on skiing holiday.
18) Experiencing a move to another place.
19) Going into an attraction at a fair.
20) Going to a national themepark