# Revictimisation and false memory formation

# Revictimisation & False Memory Formation

Bruna Calado

# Revictimisation and False Memory Formation

Bruna Calado

Cover image by: Thainá Rodrigues

Layout: Dennis Hendriks | ProefschriftMaken.nl
Printed by: ProefschriftMaken.nl

ISBN: 978-94-6423-794-8



# Revictimisation and False Memory Formation

Doctoral Thesis

To obtain the joint degree of Doctor of Philosophy
from Maastricht University and the University of Gothenburg,
on the authority of Rector Magnificus Pamela Habibovic,
according to the decision by the Board of Deans,
to be defended in public on
20th of May, 13:00 hrs in Maastricht

by
Bruna Calado

**Supervisors**
Prof. Dr. H. Otgaar, Maastricht University, Catholic University of Leuven
Dr. S. Landström, University of Gothenburg
Dr. T. J. Luke, University of Gothenburg

**Doctoral Dissertation Committee**
Prof. Dr. C. de Ruiter, Maastricht University (chair)
Prof. Dr. V. de Vogel, Maastricht University
Prof. Dr. L. Hope, Portsmouth University
Prof. Dr. A. Memon, Royal Holloway University of London
Dr. C. Lunde, University of Gothenburg

*A minha amada mãe e irmã and the rabbit I followed out of many holes.*

## Table of content

# Chapter I

## General Introduction

Calado, B., Otgaar, H., Luke, T. J., & Landström, S.

*This chapter is an adapted version of the following paper:*
Remembering what never occurred?
Children's false memories for repeated experiences.
The Inquisitive Mind Magazine, 37.

*'Memory, my dear Cecily, is the diary that we all carry about with us.'*
*'Yes, but it usually chronicles the things that have never happened, and couldn't possibly have happened.'*
*- Miss Prism and Cecily conversing about Cecily's diary in Oscar Wilde's play, The Importance of Being Earnest (1990)*

Before scientific studies demonstrated that an unexperienced event could be remembered as an autobiographical fact (e.g. Loftus & Pickrell, 1995), classic literary authors, including Oscar Wilde, had already questioned memory's trustworthiness. Arguably, in our ordinary routines, misremembering autobiographical facts (i.e. false memories) may not be an alarming matter. In fact, the subjective manner in which we remember life episodes is a constituent part of our identity (Roediger et al., 2009; Van Dyke & Alcock, 2003). The application of memory within the legal ambit, however, highlights the severity of the consequences of false remembering. More specifically, having false memories of child sexual abuse may lead to false accusations, potentially culminating in wrongful convictions (Loftus & Ketcham, 1994).

Child sexual abuse is a global problem and a widespread legal offence that occurs in a variety of societies more often than assumed (Murray et al., 2015). Within the field of psychopathology, the study of child sexual abuse began in the 1970s (Wolfe, 2007). According to Finkelhor (1984), child sexual abuse can be defined as 'sexual contact with a child that occurs under one of the three conditions: when a large age or maturational difference exists between the child and the aggressor; when the aggressor is in a position of authority over or in a care-taking relationship with the child; or when the acts are carried out against the child by using violence or trickery' (p. 101). Furthermore, the World Health Organisation (1999, p. 75) defines child sexual abuse as follows:

> The involvement of a child in sexual activity that he or she does not fully comprehend, is unable to give informed consent to, or for which the child is not developmentally prepared and cannot give consent, or that violate the laws or social taboos of society. Child sexual abuse is evidenced by this activity between a child and an adult or another child who by age or development is in a relationship of responsibility, trust or power, the activity being intended to gratify or satisfy the needs of the other person. This may include but is not limited to: the inducement or coercion of a child to engage in any unlawful sexual activity; the exploitative use of child in prostitution or other unlawful sexual practices; the exploitative use of children in pornographic performances and materials.

Within the legal system, child sexual abuse cases are difficult to prosecute for a series of reasons (Ernberg et al., 2018). Children tend not to report having been sexually

abused and, when they do, there are hardly ever witnesses to corroborate the children's claims, nor other forms of evidence such as DNA traces (Saint-Martin et al., 2007). In addition, some victims of child sexual abuse end up reporting what they suffered as a child several years after the fact, which is equally problematic to prosecute. Connolly and Read (2006) explored a number of issues related to the reliability and credibility of historical sexual abuse victims. For instance, the long-term forgetting of autobiographical information raises the question of whether or not child sexual abuse memories can always be trusted (Connolly & Read, 2002). Likely in most cases of victims recalling child sexual abuse, such recollections correspond to factual events. There is, however, a small portion of cases reported within the legal system that could correspond to fabricated memories of child sexual abuse. The latter possibility is the central theme of the current thesis. More specifically, this thesis presents four methodologically discrepant projects addressing common issues present in cases of revictimisation and potential formation of false memories. The following three subsections provide a context for the central theme of this thesis.

## The Satanic Panic Hysteria

There was an eruption of thousands of dubious incest and child sexual abuse cases during the 'satanic panic hysteria' in the 1980s in the United States. In the 1970s, a few feminist scholars attempted to raise the national awareness of child sexual abuse and incest as much more frequent features of American society than they appeared to be (Maran, 2010; also see Summit & Kryso, 1978). Proponents of this movement argued that revictimisation was a common characteristic of child sexual abuse. In fact, the literature indicates that a considerable number of children who survive sexual abuse are likely to be abused again in the future (Walker et al., 2019). In the 1980s to 1990s, the awareness of the ordinariness of repeated child sexual abuse reached its peak, bringing feelings of guilt and distress to families (Yuhas, 2021).

The understanding of the incidence levels of child sexual abuse and incest shaped some psychotherapy methods, as more alleged sexual abuse victims sought professional help. Consequently, the number of self-help books to aid historical abuse survivors who could not recall being sexually abused also increased. The bestseller *The Courage to Heal* by Ellen Bass and Laura Davis, which was published in 1988, offered clear instructions on how to recover repressed memories of child sexual abuse. During the 1980s and 1990s, the sexual abuse crisis that had been established in the United States brought the classic psychoanalytic concept of repression back into popularity (e.g. Cohen, 1985; Davis & Schatzow, 1987). Repression is considered to be a defence mechanism that takes place to protect a person's psyche from traumatic experiences. When a person experiences a highly distressing event, the memory of the experience is removed from the person's consciousness and placed in their unconscious (Freud, 1961).

In the 1980s and 1990s, many psychotherapists in the United States used the concept of repression to actively aid their patients to recall child sexual abuse memories that could have been repressed (see Andrews et al., 1995). Furthermore, psychotherapists offered specific treatments for incest survivors that aimed to bring repressed memories of child sexual abuse to the surface (e.g. Meiselman, 1990; also see Coker, 1990). Incest survivor therapy and books shared a common instruction, recommending people not to dismiss any thoughts related to child sexual abuse.

According to methods that focused on recovering repressed memories, if anyone believed that they could have been abused as a child, there was a high likelihood that abuse had in fact happened to them (Lindsay & Read, 1994). Moreover, various psychological disorders were associated with child sexual abuse. For instance, therapists considered depression during adulthood to be one of the main indicators of child sexual abuse (Browne & Finkelhor, 1986). However, the consequences experienced by a child sexual abuse victim are varied and subjective (Kennedy & Prock, 2018). The problem with perceiving depression as a direct symptom of child sexual abuse is that it presupposes that every victim of incest or other forms of child sexual abuse would invariably develop depression, which is not necessarily the case (Putnam, 2003). Hence, clinically depressed patients who had not been sexually abused as a child could become convinced that they were among the many child sexual abuse survivors emerging during that period.

Associating various multi-causal psychological conditions with child sexual abuse fomented an atmosphere of panic, in which people feared not only that children could be repeatedly sexually abused at school, but also that parents themselves could have been abused as children multiple times without remembering it (Hughes, 2017). Thus, the social movement to raise awareness of child sexual abuse can be interpreted both positively and negatively. On the positive side, it was obvious that many child sexual abuse survivors were able to speak openly about their traumatic experience. On the negative side, this burgeoning awareness created a form of mass social hysteria that could have facilitated false reports and, in a worst-case scenario, false memories (Lief & Fetkewicz, 1995).

The McMartin Preschool case, which will be described below, is an example of a case in which false reports and/or false memories were likely fostered (Schreiber et al., 2006). The State of California investigated the McMartin Preschool case for over 7 years but was never able to prove the children's reports of being repeatedly sexually abused. The McMartin children also described supernatural phenomena and secret rooms that were never found in the school. Because the children underwent an enormous amount of suggestive interviewing, the literature argues that the statements provided by the McMartin children were rendered unreliable and could be the product of false memories (Garven et al., 1998). Before describing the McMartin Preschool case, I will first present a self-reported case of false memories of revictimisation and child sexual abuse.

## The Meredith Maran Case

During the period of the sexual abuse panic, in the 1980s, Meredith Maran, an award-winning American journalist, dwelled on the possibility that she had been repeatedly sexually abused by her father as a child. Maran and her father used to be very close when she was a child; however, their relationship had declined drastically since her adolescence. As an adult, Maran had always looked for answers that could address the radical change in their relationship.

During this streak of sexual abuse panic in the 1980s, tens of thousands of Americans believed that they had repressed hideous memories of incestuous child sexual abuse (Hughes, 2017). In that period, Maran was assisting another journalist in writing a book about incest survivors in the United States. After some time working with such cases, Maran read *The Courage to Heal* (Bass & Davis, 1988), which at first led her to believe that her father could have sexually abused her as a child.

As Maran began to dream about incest episodes with her father, she decided to seek specialised therapy for incest survivors. Like many others, Maran's therapeutic sessions were specifically targeted at aiding patients to recover repressed memories of incest. Hence, the therapist would, for instance, recommend that the patients ease their mind into the idea that they had been sexually abused as a child by their father or other family members. Therapists would use tools such as hypnosis, as well as guiding patients' imagination to recreate the abuse scenarios (see Lindsay & Read, 1994; Maran, 2010).

Throughout Maran's therapeutic treatment, her dreams increased which, contributed to her belief that her father had in fact abused her when she was a child. In the course of time, Maran formed rich autobiographical memories of being recurrently sexually abused by her father despite claiming to never be fully confident that her memories were authentic and not a product of her imagination. After some time, Maran informed her family about the discovered outcome of her therapeutic process. Consequently, Maran's father was cast out from the family conviviality.

Years later, Maran was confronted with several cases of false memory, which reached the first page of newspapers and were broadcasted on television. After reading many scientific papers addressing the issue of repressed memories, she realised that her memories had been a product of suggestive therapeutic tools and a highly socially influenceable atmosphere. By accepting that she could have been one of the many reported victims of that time, a path was paved for Maran to create memories that could fill in the gap of the broken father and daughter relationship in her story (Maran, 2010). She interrupted her treatment to recover her 'repressed' memories and ceased attending support groups for incest survivors. As she distanced herself from that universe, her recurrent dreams of incest disappeared. Eventually, she reconnected with father, who had always denied her accusations.

## The McMartin Preschool Case

Contemporaneous to Maran's case, the McMartin Preschool case involved multiple child sexual abuse allegations that were later deemed to be false. The McMartin Preschool case started with one child's report, and then grew to over a hundred reports from children enrolled that year and former students of the same preschool in Manhattan Beach, California. Many of these children declared that they had been victims of sexual abuse by the McMartin Preschool staff (Wyatt, 2002). Children from other schools charged their teachers as well, bringing the number of allegedly abused children during the investigation to over 1400 across the State of California (Haberman, 2014). The apparent consistency between the reports, the level of detail presented, and the high number of children involved led practitioners (i.e. clinical psychologists, social workers, legal professionals and police) to believe that the allegations were credible. However, the interviewing techniques used by the professionals likely contaminated the children's testimonies.

The case started with a statement from the parent of an allegedly abused child in 1983. Without further investigation, the local police sent a letter to the families of the 200 students who were enrolled at the McMartin Preschool at that time. In the letter, the police warned parents about the possibility that their children could also have been sexually abused at school. Hence, the police team urged parents to question their children about being sexually abused by the McMartin Preschool staff. Problematically, research shows that parents may unintentionally question their children in a suggestive manner that may lead to false reports, with the potential consequence of jeopardising the children's real memories (Korkman et al., 2014; Thompson et al., 1997).

Moreover, the interviews with children that were conducted by social workers were rather suggestive. When analysing the McMartin Preschool case, memory scholars argued that the children had been led to say that repeated occasions of sexual abuse had taken place (Garvin et al., 1998). There were several interviewing situations in which the children denied to the social workers that they had been sexually abused; yet the social workers gave the children negative feedback, attesting that their answers were not sufficient. A short example of such an interview can be seen here:

| | |
|---|---|
| Interviewer: | Can you remember the naked pictures? |
| Child: | (Shakes head 'no') |
| Interviewer: | Can't remember that part? |
| Child: | (Shakes head 'no') |
| Interviewer: | Why don't you think about that for a while, okay? Your memory might come back to you. (Garvin et al., 1998; p. 28) |

In addition, when the children gave a satisfying answer, they were positively

reinforced by the social worker stating, for example, 'Well done, I knew you were smart enough to remember.' (Garven et al., 1998, p. 27). These and other types of suggestive prompts have been shown to increase the chance of false reports (Garven et al., 1998; Schreiber et al., 2006). Many children in the McMartin Preschool case went along with interviewers' suggestions, causing their testimonies to contain some bizarre elements, such as experiencing satanic rituals, seeing witches fly on brooms, and children being flushed down toilets. False memory researchers suggested that the (repetitive) interviewing process in this case might have led the children to agree with the proposed suggestions and possibly form false memories (Brainerd & Reyna, 2002; Garven et al., 1998).

In addition, the liberty with which the McMartin children were able to exchange information about the alleged occurrences of abuse with one another could have been an important source of misinformation. The interviewers misinformed the children using supposed information they had collected from the children's classmates (Schreiber et al., 2001). In such circumstances, it might have been the case that the McMartin children were exposed to memory conformity situations. Social contagion of memory, or memory conformity, is a phenomenon that occurs when people influence one another's memories (Roediger et al., 2001). To be specific, when collectively experiencing an event, each person holds subjective memories that do not necessarily contain the same information as the memories of the rest of the group. Memory conformity happens when a member of the group incorporates memories belonging to their peers as their own. Memory conformity takes place due to both social influence and source monitoring errors (Gabbert et al., 2004).

The McMartin case ran from 1983 to 1990, and the State of California spent over $16 million USD on a 7-year investigation (Wyatt, 2002). Because of the lack of physical evidence corroborating the statements given by the children, the case was closed without any convictions. The considerable number of cases with a similar nature to this one (but varying in terms of severity, the number of alleged victims etc.) stresses that the McMartin Preschool is not an exceptional case (Otgaar et al., 2017). Assuming that the children's allegations are in fact false, serious questions can be raised about how such detailed statements could be produced. One possibility is that the children merely acquiesced to the investigators' suggestions. Another, even more disturbing, possibility is that some of the children came to truly believe and remember that they had been sexually abused, despite abuse never actually having occurred.

### *Children's Suggestibility in Forensic Contexts*

The McMartin Preschool case is an extreme illustration of how suggestive interviewing techniques can compromise children's ability to provide a trustworthy statement. In the McMartin Preschool case, the social workers who interviewed the children presumably believed that the McMartin Preschool staff had sexually abused the children. Such a

belief is likely to have shaped the social workers' questions in a manner that led the children to disclose the information that the social workers expected to hear, rather than the truth. Arguably, the social workers influenced the children's narratives (Garvin et al., 1998).

In terms of children's developmental characteristics, there are some contexts that can position children as suggestible subjects. For instance, children are naturally inclined to please authority figures. The social influence exercised by adults over children can lead children to provide adults with whatever information the children anticipate the adults want to hear, even when such information is incorrect (Ceci et al., 1997). Moreover, compared with older age groups, young children have a limited memory-monitoring ability (Ceci & Bruck, 1995). Thus, if children misattribute the source of a memory, it is not difficult for them to confound a suggested event with an experienced event.

Besides source monitoring difficulties, children's encoding can be weakened when their memory is already deteriorated due to time or a lack of attention. Hence, children's weak representations might be more easily overwritten by suggestions compared with children's vivid representations (Brainerd, 1990; Bruck & Ceci, 1999). Scientists have attempted to develop structured interview protocols that are appropriate to children's particular developmental characteristics. Child interviewing protocols aim to optimise children's performance and protect children from interviewers' biases.

In structured interviewing protocols, such as the National Institute of Child Health and Human Development (NICHD) protocol (Lamb et al., 2007), scientists orient practitioners to start interviewing children with open-ended questions also called invitations. That is, when asked an open-ended question such as 'Tell me what happened', children are preserved from receiving an input from the interviewer that could potentially overwrite their actual memory. Furthermore, such interviewing protocols account for children's optimal length for each interviewing session, based on children's developmental characteristics. In the McMartin Preschool case, the children were repeatedly interviewed for over 1 hr, without breaks. Long interviews might compromise pre-schoolers' attention due to exhaustion (Cordisco Steele & NCAC, 2015). Moreover, an analysis of the interviews' transcripts shows that the interviewers spoke far more than the children, which indicates a lack of use of open-ended questions, as well as the use of suggestive techniques such as repetition of the questions and reinforcement (Schreiber et al., 2006; also see Poole et al., 2014).

Children also have limited symbolic representational ability (Deloache & Smith, 1999). Hence, if a child is provided with props such as drawings and anatomic dolls, the child can easily confound what they are visualising with the memories they are trying to convey in their statement (Bruck et al., 2000). For example, suppose a child is asked to describe the colour of the car she saw abducting her friend. In order to remind the child of what happened, the interviewer offers her a car toy to play with. Considering that the car toy will definitely have a colour, it is possible that the child being interviewed will

name the colour of the car toy she is seeing at that very moment, rather than the colour of the car she saw abducting her friend.

The McMartin children played with a variety of toys and received different visual stimuli, such as puppet theatre, during the investigative interviews of the case. These features might have inflated children's narrative of unproven child sexual abuse. Just like adults, children's encoding, storage and retrieval of episodic information are influenced by pre-existing knowledge, motivation and beliefs (Ornstein et al., 1998). Because the interviewing method used in the McMartin Preschool case disregarded child suggestibility, scientists agree that what happened in that case is the antithesis of good practice in child interviewing (Schreiber et al., 2006).

## Revictimisation and False Memory Formation

Drawing from both Meredith Maran and the McMartin Preschool case, the work presented in this thesis aims to address different dimensions pertaining to legal cases of revictimisation and potential formation of false memories. Via different methodologies, I looked into three particular aspects observed in cases akin to the McMartin Preschool case, namely false memory formation as a product of co-witness interaction, credibility attested to sexual abuse revictimisation statements, and formation of false memories of repeated false autobiographical events. Specifically, this thesis comprises: (1) a legal psychological analysis of a current potential case of false memories of repeated events; (2) a lab study to investigate how children and adolescents perform as witnesses within a memory conformity scenario; (3) an online experiment exploring the perceived credibility of sexual abuse victims when reporting single-occurrence abuse versus repeated-occurrence abuse; and, finally, (4) a laboratory experiment addressing the proof of concept that it is possible to implant false memories of repeated events under laboratory conditions. Together, these studies provide relevant practical implications within the theme of revictimisation and false memories. The subsequent subsections delineate relevant theories and research that compose the background of the research I developed in this thesis.

## (False) Memory Frameworks

What happened to Maran and, potentially, to some of the children at the McMartin Preschool is an illustration of false memory cases of repeated events and their potentially harmful implications in the legal field. Assuming that false memories of repeated autobiographical events can in fact take place under specific conditions, (false) memory theories provide tentative predictions of why and how such memory phenomena can occur. In this section, I briefly present the contributions of three (false) memory frameworks that are relevant to the matter of false memory implantation of

repeated events: *schema theory, fuzzy trace theory,* and *associative activation theory*.

## Schema Theory

More generally situated within the classic memory framework, *schema theory* explains the relationship between the general knowledge an individual holds about how repeated experiences are structured within a person's memory, and an individual's ability to dissociate specific instances from those experiences (see Bartlett, 1932; Friedman, 1979; Schank & Abelson, 1977). Bartlett (1932) stated that schemas are structures of general knowledge that influence memory retrieval and impact memory reconstruction.

As an individual repeats a certain event or task, structured representations of the occurrences are stored in the individual's memory. Hence, a schema provides an organisational setting that assists with the memory and perception of similar experiences. For example, after a year of practice, a chess player forms structured knowledge about how to behave in a chess match, in terms of both social norms and game rules. Schemas are formed for any repeated experience; hence, we all have schemas for parts of our lives. Through schemas, it is possible to encode and store pertinent information and efficiently combine new information with already existing schemas.

Schemas also influence the allocation of attentional and encoding processes (Anderson, 1983), which suggests that schemas have an associative nature that combines a variety of elements belonging to a given experience (Ghosh & Gilboa, 2014). The adaptive and flexible characteristics of schemas allow memory reconstruction. More specifically, the information attached to a certain episode might be allocated to a different episode when there are similarities between both episodes – a relevant characteristic for false memory (of repeated events) formation. Furthermore, an individual may intuitively understand that repeated experiences should foster a schematic knowledge concerning that event. Having such schematic knowledge about a reoccurring event should consequently raise a sentiment of familiarity every time the individual experiences that event.

## Fuzzy Trace Theory

Specifically looking into false memory issues, the *fuzzy trace theory* explains that, when experiencing an event, two independent memory traces are formed: gist and verbatim (Brainerd et al., 2008; Reyna & Brainerd, 1995). Gist traces involve the general meaning of an event (e.g. playing chess) and can be retrieved even after a long delay (Brainerd & Reyna, 2012). Verbatim traces include the precise details of an experience (e.g. the specific chess moves used in a match) and fade rapidly.

The deterioration of verbatim traces over time is the reason why people can have difficulty retrieving specific details about an event as time goes by. When they forget

verbatim traces about an event, people mainly rely on the general meaning of what happened, which may enhance false memory proneness. Being uncertain about the details of an event (verbatim traces) may contribute to the creation of wrong associations of details in the memory of an event (e.g. believing one used a French Defence to open a chess match, when that move was actually used by one's opponent). Fuzzy trace theory suggests that the memory's characteristic of deteriorating over time, particularly in terms of details associated with the experience of an event, is directly related to the formation of false memories.

### *Associative Activation Theory*

*Associative activation theory* is another framework that is intended to explain false memory formation. Associative activation theory explains that the semantic representations of life experiences are stored in the memory and can be activated by corresponding concepts (Howe et al., 2009). This theory stems from 'spreading activation models' (see Anderson, 1983), which infer that, when processed, one word can activate corresponding concept nodes within a person's knowledge base, so the activation may scatter to surrounding concept nodes (Collins & Loftus, 1975; Landauer & Dumais, 1997). As a result, not all concepts will be correctly evoked in such association processes, and incorrectly activated concept nodes may be erroneously remembered as an actual feature of a life experience, thereby generating a false memory (Otgaar et al., 2018).

## False Memory Paradigms Relevant to this Thesis

Several decades of advances in (false) memory framework directly contribute with methods invented by researchers to assess false memory formation in different contexts. Below, I describe three memory paradigms (methods) designed to elicit different types of false memory used in two chapters of this thesis.

### *The DRM Paradigm*

Unlike false memories that arise from external misinformation, spontaneous false memories are a product of endogenous memory mechanisms such as spreading activation or a reliance on schematic knowledge (Brainerd et al., 2008). A solid method to induce spontaneous false memories is the Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). The DRM paradigm procedure consists of eliciting the 'critical lure' of a thematic cluster of words via word lists that are given to participants to memorise. The critical lure is a word that is not present in the word list, but is evoked by the other related words in the list. For example, the critical lure of 'swallow, chew, food, nourishment, taste' might be the word 'eat'.

Researchers can ask participants to report the words they remember from the word list in two ways. In the first way, researchers invite participants to freely recall the words

they viewed in the word list(s). In the second way, researchers present the participants with a longer list of words from which the participants can recognise the words they saw previously. Studies using the DRM paradigm consistently find that participants report remembering to have read the critical lure in the word list, even though the critical lure was absent from the word list (Huff et al., 2015).

Both the fuzzy trace theory (Brainerd et al., 2008) and the associative activation theory (Howe et al., 2009), previously described in this chapter, explain the formation of spontaneous false memories. To be specific, the fuzzy trace theory posits that when an item (e.g. a critical lure) that does not belong in the original memory (e.g. a word list) is associated with the gist of that memory, the item can be falsely remembered as part of the memory. The erroneous association of the item with the gist of the memory particularly occurs due to a lack of verbatim traces of the memory in question. When verbatim traces are absent, there is no contradictory information to prevent unoriginal items from being wrongly associated with a memory.

The associative activation theory explains the formation of spontaneous false memories as occurring through an individual's knowledge base and the automatic activation of information. Thus, an element can activate several corresponding concepts that are part of an individual's knowledge base, but that were not originally part of a memory, by incorporating them through association with the memory in question. The associative activation theory helps to explain why children are less prone than adults to form spontaneous false memories when freely recalling a past episode – an effect known as developmental reversal. Since children's knowledge base naturally contains less information than adults' knowledge base, children have fewer concepts to erroneously associate with a given past experience.

***Memory Conformity Paradigm***

While children might be reliable witnesses when freely recalling what they saw, children can also produce untrustworthy statements when placed in a context of social pressure (e.g. the McMartin Preschool case; McGuire et al., 2015). In cases akin to the McMartin Preschool case, victims may be invited to report their memories after sharing their impressions with other victims. When individuals share information with one another before reporting what they remember about the crime of which they were victims, there is a possibility of one victim contaminating the other victim's memories with her own and *vice versa*. In a similar case to the McMartin Preschool case, for instance, assuming children's report contained factual information, the quality of the information could be deemed dubious as the children exchange information before providing their reports (Gabbert et al., 2012).

To assess whether co-witness interaction can contaminate witnesses reports with foreign information or memories, scientists rely on the memory conformity paradigm. As the name suggests, the memory conformity paradigm investigates false memory

formation by the propensity of individuals to conform with a peer's version of a given event to the point of incorporating such version as their one (Wright et al., 2000). The memory conformity paradigm occurs when a participant witnesses a scene and is later allowed to chat with a confederate who has observed the same scene. Subsequently, the participant reports what she can remember from the scene she observed (Gabbert et al., 2003). In this paradigm, the confederate misinforms the participant regarding some observed elements of the scene in question. Studies utilising this paradigm consistently find that, despite being asked to report what they can remember themselves about the scene they observed, participants tend to report the misinformation they received from their confederate. Several studies on the topic of memory conformity advanced the investigation on relevant factors for this topic.

For instance, Allan and Gabbert (2008) investigated in two experiments the nature of social influence and social influence permanence in a co-witness condition. Their first experiment (*N=25)* showed that an individual's immediate conformity is modulated by the social proximity to the confederate as well as by the individual's confidence in the trustworthiness of the confederate's version of what they witnessed. Furthermore, their second experiment (*N*=22) showed that the misinformation given by the confederate about the witnessed event remained in individuals' memory a day later. The results of both experiments suggest that individuals tend to not doubt the veracity of peers' statements when there is no suspicion of intentional misinformation. Such findings are useful when trying to understand the interpersonal influence the McMartin Preschool case children (and children in similar cases) had over each other regarding potential memory distortion.

***False Memory Implantation Paradigm***

The implantation method – also known as the 'lost-in-the-mall' paradigm – mimics the situation that has been found in dubious child sexual abuse cases from the satanic panic period. Loftus and Pickrell (1995) conducted the first experiment demonstrating the formation of false memories of entire autobiographical events. They used a sample of 24 participants. In Loftus and Pickrell's (1995) experiment, the researcher prepared one booklet containing four short narratives about the participants' childhood, in which three were true autobiographical narratives that had been confirmed by the participants' parents, and one was a false narrative produced by the researchers.

The researchers manufactured the false narrative, which involved getting lost in a big shopping mall around the age of 5. The details of the story included crying a lot, being found by an elderly person and being reunited with the family. The researchers made sure to check with the participants' families to determine whether the participants had experienced the manufactured narrative. The participants were interviewed about their childhood narratives (including the false narrative) on two separate occasions. Of the total sample, 25% eventually reported having remembered being lost in a shopping

mall and were confident that such an event had happened to them.

After Loftus and Pickrell's (1995) experiment, memory researchers replicated the experimental design with a mixture of true and false narratives, and examined different factors that could affect the formation of implanted false memories. On average, over eight published studies, approximately 30% of the participants indicated that they remembered the false event (Scoboria et al., 2017). For example, research has shown that negative false events are easier to implant in children than neutral ones.

Otgaar and colleagues (2009) had 7-year-old children ($N$=72) listen to three narratives, one of which was false. The false narrative could be either neutral (moving to another classroom) or negative (being accused of cheating by the teacher). Each child went through two interview sessions, with a gap of a week between them. Besides showing that some children formed rich false memories, this study showed that more participants fell prey to the implantation of the negative false event compared with the neutral false event.

Other studies demonstrated that the plausibility of a false event was an important catalyst for false memory implantation in adult participants. Pezdek and colleagues (1997) showed that, out of 51 participants, 22% ($n$=11) formed false memories of a plausible event, whereas only 6% ($n$=3) formed false memories of an implausible event. Despite radically improving the understanding of memory's fragility and the false memory formation of autobiographical events, false memory implantation studies have focused on studying the implantation of single-occurrence events (Brewin & Andrews, 2017). Attempting to implant an event that took place a single time is not necessarily representative of child sexual abuse cases, since revictimisation is a common feature in such cases (Blizard & Shaw, 2019).

## False Memories of Repeated Events

Research that is relevant to the theme of false memory formation for entire repeated autobiographical events includes investigations of the mechanisms of false memory formation for truly experienced repeated events. When an event is repeatedly experienced, people form a *script* (as per schema theory) for this event, which is a mental representation of the general sequence of elements that are part of the repeated experience (Connolly & Price, 2006). When scripts are formed, the specific details of an event become harder to retrieve, and people usually rely more on the gist (as per fuzzy trace theory) of the whole sequence of events than on specific memories.

Therefore, when people are exposed to repeated experiences, it is difficult for them to distinguish the verbatim traces (as per fuzzy trace theory) of one specific experience from the other experiences of a repeated event. In other words, individuals find it difficult to identify specific incidents of a repeated experience. As a result, people tend to rely on the meaning of the event as a whole (gist traces) and forget about important pieces

of the different occasions (verbatim traces). If, for instance, an individual is questioned about an event in which the verbatim traces for the corresponding memory have already deteriorated, the individual may automatically fill in the gaps in her memory with similar information. Therefore, a lack of verbatim traces and a reliance on gist traces may fuel the formation of false memories (Brainerd & Reyna, 2012).

As mentioned above, studies have compared the formation of false memories of events that happened once (i.e. single events) versus those of events that happened more than once (i.e. repeated events) (Connolly & Price, 2006; Connolly et al., 2008[b], 2016; Price et al., 2006). Children can form false memories from single and repeated events. Connolly and Price (2006) conducted a study in which they hypothesised that the association between different details of the same repeated experience could help to explain these discrepant findings.

In their study, preschool (4- to 5-year-old) and first grade (6- to 7-year-old) children were assigned to either a single play session or four repeated play sessions for a total of eight sessions. For each of the eight play sessions, in both the single and the repeated groups, two critical items were presented to the children, for a total of 16 critical details. In one of the eight possible play sessions, the children were instructed to reach for water (critical detail 1) and pretend to be a dog (critical detail 2). Two weeks later, the children received misinformation regarding half of the critical details. On the following day, the researcher asked the children to freely remember details about one specific session; subsequently, the children received cues to help them remember that session.

The 6- to 7-year-olds in the repeated event group were more suggestible to the creation of false memories than those in the single-event group, particularly for high-association details. Similarly, the 4- to 5-year-olds in the repeated event group were more suggestible to the formation of false memories than the single-event participants, but only for low-association details. This study showed that the memories of repeated experiences can be quite fragile and are susceptible to false memory formation.

Studies on the false memory formation that arises from repeated life experiences can provide a notion of how false memories of repeated experiences are produced. Some of these studies have demonstrated that children who experience repeated events are more prone to produce false memories than children who experience a single event (Connolly & Lindsay, 2001; Price & Connolly, 2004). For instance, in another study by Price and Connolly (2004), 90 4- to 5-year-olds participated in either one or four play sessions. Subsequently, when the children were interviewed about the play sessions in a suggestive way, half of the recalled details were wrong. The children assigned to the repeated event group reported more mistakes about their play experience than those assigned to the single-event group. From a broader perspective, these studies show that if a similar life event is repeated several times, it becomes difficult for people to accurately remember the specific details of each instance of the event in question.

## Credibility Assessment for Repeated Experiences

As previously stated, child sexual abuse has high incidence of revictimisation. Moreover, in child sexual abuse cases, it is common that no physical evidence is found to corroborate the victims' and witnesses' statements (Weinsheimer et al., 2017). Hence, credibility assessment is a common task in the investigation of child sexual abuse cases. Such characteristics are likely to lower the victims' perceived credibility (Connolly et al., 2008[a]).

The perceived credibility of repeated events has practical implications for eyewitness testimony evaluation. Some studies have focused on assessing children's perceived credibility when reporting a unique experience versus repeated occurrences of the same thematic experience (see Connolly et al., 2008[b]; Connolly & Lavoie, 2015). These studies consistently found that children were more credible when reporting a single-occurrence experience than when reporting a repeated experience. These studies did not use sexual abuse statements *per se*; however, it is expected that such results can be generalised to this ambit, as well as to other legal contexts.

## The Thesis Outline

Chapter II reports on the Jakarta International School (JIS) case. This recent day care case of child sexual abuse features three children who were subjected to extensive suggestions from different sources over time, and who eventually came to report being repeatedly sexually abused by the school staff. The accused staff members were prosecuted and sentenced to jail based on doubtful allegations of repeated episodes of sexual abuse against the children. The repercussions of the JIS case seemed to contradict the literature on credibility assessment for repeated versus single experiences. That is, when people recount repeated episodes of the same experience, they are deemed to be less credible than people recounting a unique experience. The literature on the topic of the credibility assessment of repeated experiences in regard to sexual abuse allegations is limited by the fact that the studies in the literature did not use sexual abuse statements for their participants to assess.

In the JIS case, the supposed experiences of sexual abuse were widespread, with allegations of abuse by one child to another, and by adults to children. Chapter III presents a laboratory experiment in which I assessed children's and adolescents' propensity for false memory formation in a co-witness situation. Furthermore, I investigated developmental reversal trends for our participants through the memory conformity paradigm and the DRM paradigm to assess whether children could perform better than adolescents under certain circumstances. This study is relevant for understanding the consequences of allowing children to be exposed to social pressure and misinformation before providing a legal statement, as occurred in the McMartin Preschool case.

Chapter IV presents an online study in which I used vignettes based on actual statements of sexual abuse allegations to be assessed by participants in terms of credibility. I was mainly interested in learning whether it made a difference for the participants if the victims of sexual abuse recounted multiple incidents of sexual abuse by the same perpetrator or a unique incident of sexual abuse. Furthermore, I wanted to assess whether the participants would attribute different levels of credibility to children reporting current sexual abuse in comparison with adults reporting current sexual abuse and adults reporting historical sexual abuse.

In many potential cases of false memory of sexual abuse, including the JIS case, children have reported repeated instances of abuse. The literature reports a variety of aspects that are relevant for autobiographical memory implantation. However, to the best of my knowledge, there is no scientific evidence showing that false memories of repeated events can be implanted under laboratory conditions. Chapter V reports a laboratory experiment that demonstrates that false memories of repeated events can be implanted. Finally, Chapter VI discusses the key findings gathered throughout the research that comprises this thesis. It further presents practical implications of the current work, as well as limitations and suggestions for future research.

# Chapter II

## Suggestive Questioning and Allegations of a Magic Stone: An Analysis of the Jakarta International School Case from the Perspective of Legal Psychology

Calado, B., Sumampouw, N., Otgaar, H., Luke, T. J., Landström, S., & London, K.

*This chapter corresponds to the following paper:*
Suggestive questioning and allegations of a magic stone: An analysis of the Jakarta International School case from the perspective of legal psychology.
Forensic Science International: Mind & Law. (under review)

**Abstract**

In 2014, allegations of a child sexual abuse ring emerged at the Jakarta International School (JIS) in Indonesia. The case featured three kindergarten children who ultimately reported having been repeatedly sexually abused by different members of the staff during the course of their regular school days. Six janitors working for the International Service System (ISS) placed at the JIS and two teachers working for the JIS were sentenced to several years in prison for the alleged sexual assaults. Among the cleaners, the only accused female cleaner was sentenced to 7 years in prison, while the five accused male cleaners each received sentences of 8 years. Both teachers received sentences of 11 years in prison. The investigation records and the court reports of the JIS case indicate that the three children were heavily exposed to suggestive interviewing techniques by their parents, the police and psychotherapists. According to the literature, repeated suggestive interviewing may foment false reports or false memory production in children. Furthermore, to the best of our knowledge, the children repeatedly underwent invasive medical tests for child sexual abuse that constantly gave negative results. In our opinion, the evolution and content of the children's allegations of revictimisation share commonalities with some of the large-scale unproven child sexual abuse cases that have been documented in Westernised countries. In their accounts, the children in the JIS case included implausible details that were elicited over the course of a lengthy process using highly suggestive methods. We discuss the aspects of the JIS case in light of the scientific literature on memory and children's suggestibility, and argue that the methods used to elicit the children's reports in the JIS case parallel techniques that produce false reports or false memories. Recommendations about forensic child interviewing and criminal investigation procedures are presented.

***Keywords:*** JIS, day care, sexual abuse, child suggestibility, false memory

## Opening

During the wave of 'day care sex abuse hysteria' or 'Satanic Panic' in the 1980s and 1990s, several unfounded cases featured hundreds of children in the United States and other Western countries (Loftus & Ketcham, 1994). In these cases, the children typically claimed to be repeatedly sexually abused by the staff of their school (for reviews, see Nathan & Snedeker, 2001; Young, 2008; see also Goodman et al., 1997)[1]. Common threads run through the unproven day care sexual abuse cases. In such cases, the children's allegations – which often included unrealistic elements – emerged gradually in the context of highly suggestive interviewing techniques. In other words, the children's disclosure of abuse was not spontaneous, but may have been a product of repeated exposure to social influence and suggestive questions (Garven et al., 1998).

In these cases, various circumstances seemed to convince the parents that their children had suffered sexual abuse (e.g. Rosenthal, 1995). The circumstances that make parents believe that their child may have been sexually abused are often related to the children's behaviour. Initially, when questioned by their parents, children often and repeatedly deny abuse (see Rosenthal, 1995; Garven et al., 1998). Counterintuitively, adults tend to perceive children's denials as consistent with abuse. This perception leads parents to persevere with high-pressure questioning of the children, with the intention of protecting the children. Consequently, the children in these cases may eventually come to make claims that are consistent with the adults' beliefs (see Rosenthal, 1995; Garven et al., 1998).

In response to the many unproven cases of child sexual abuse revictimisation (see Table 1 for examples of cases reported in the scientific literature), psychologists have conducted research to guide developmentally sensitive forensic interviews. Numerous evidence-based forensic interview protocols are available worldwide (Holliday, 2003; Lamb & Brown, 2006; Lamb et al., 2007). As these interview protocols make it possible to collect more reliable statements from children, some scholars may believe that unproven sex abuse allegations like these are a relic of the past. However, such cases still occur. Here, we describe and analyse a recent day care abuse case that started in 2014 in Jakarta, Indonesia. In this case, two schoolteachers and six janitors of the Jakarta International School (JIS) were convicted based on unrealistic statements from three children. Due to the suggestive manner in which the children's reports were elicited (without the use of validated interviewing protocols) and the lack of physical evidence to corroborate the children's statements, it is our opinion that the children's statements in the JIS case were rendered unreliable.

1 Both then and now, child maltreatment is widely agreed upon to be a major international public health problem. However, the cases that arose in the context of the mass hysteria share common features and are qualitatively different from high-probability cases of maltreatment, as is later reviewed in this paper.

**Table 1.** *Examples of Unproven Cases of Child Sexual Abuse Revictimisation Documented in the Scientific Literature*

| Name | Year | Country | Summary | References |
|---|---|---|---|---|
| *McMartin Preschool* | 1983 | United States | Over 100 children who attended the McMartin Preschool accused their teachers of sexually abusing them. Children described events such as flying on broomsticks to a secret farm, secret tunnels and seeing other children being flushed down toilets. Children were interviewed for an extensive period with highly suggestive techniques. The investigation lasted for 7 years, and no corroborative evidence was ever found. By the end of the case, the accused were found not guilty. | Campo et al., 2006; Schreiber et al., 2006 |
| *Kelly Michaels/ Wee Care Nursery* | 1987 | United States | Children accused teacher Kelly Michaels of forcing them to lick peanut butter off her genitals and of penetrating them with knives. The teacher was convicted on 115 counts of sexual offenses involving 20 children. After 5 years in prison, she was released, as the court realised that the interviews conducted with the children in this case were highly improper and suggestive. | Campo et al., 2006; Rosenthal, 1995 |
| *Oude Pekela* | 1987 | The Netherlands | A 4-year-old boy with anal bleeding disclosed during therapy to have been sexually abused. Eventually, nearly 100 children in the village were questioned about possible satanic abuse. Some of the statements described objects being inserted into children's genitalia and rectums, faeces and urine ingestion, and helping with murders. Physical and other corroborative evidence was never found, and the case was eventually dismissed. | Jonker & Jonker-Bakker, 1991; Rossen, 1992 |
| *Galileo Elementary School* | 2009 | The Netherlands | The father of one child charged two teachers with sexual abuse allegations. The day after the charges were filed, the school messaged all parents of enrolled students. Subsequently, 20 children reported having also been abused by the same teachers. Unrealistic elements were present in the children's statements, such as having a piece of their genitals cut off and seeing crocodiles at the teachers' houses. The teachers were not prosecuted, as there was not enough evidence. | Otgaar et al., 2016 |

## Chapter Structure

First, we present an overview of the JIS case, including a timeline of the main events. In the subsequent section, we explain the origin of the case, followed by forensic and developmental findings illustrating how professionals can validly assess cases with child witnesses. In the next section, we present the evolution of the allegations in the JIS case in light of cascade bias effect theory. Subsequently, to

contextualise our analysis of the JIS case, we briefly review, in separate sections, the scientific literature on (1) child suggestibility, false reports, and false memory; (2) line-up and facial recognition; (3) contemporary best practice for child forensic interviews; and (4) alternative scenario investigation. We conclude by recommending appropriate practices that are relevant for cases in which children's testimony is the only evidence available. We believe the JIS case to be an illustration of how a lack of knowledge on child investigative interviewing and eyewitness reliability in potential cases of false memory of repeated events may jeopardise the trustworthiness of children's statements, resulting in potential miscarriages of justice.



## The Jakarta International School Case

The information presented in this report comes from the Indonesian Supreme Court open-access database, which includes police reports (*Bantleman vs. Indonesia*, 2015; 2017), as well as police case records[2] and media sources. We have protected the identities of the families involved in this case by referring to them with anonymous labels. Chronologically, the first child and parents to make sexual abuse allegations against the JIS staff are represented in this report as Child 1, Mother 1, and Father 1; the second family as Child 2, Mother 2, and Father 2; and the third family as Child 3, Mother 3, and Father 3.

The last author of the current paper served as an expert witness in this case, which has since been legally resolved. The current report reflects the authors' opinions, derived from scientific evidence, based on the evidence gathered by the defence, as well as from the publicly available official reports from the Indonesian courts. The main focuses of the current report are child interviewing and suggestibility. However, due to the complexity of the JIS case, this paper also discusses legal psychological aspects that could have interfered with the investigation of the case as a whole, such as the cascade bias effect. The aim of this paper is to contribute to the literature on unproven day care sexual abuse cases with a timely example.

The JIS case began in March of 2014 and involved two 6-year-old and one 7-year-old kindergarteners, who eventually claimed that six outsourced janitors and two schoolteachers had serially abused them in an organised sex abuse ring at the school. The case materials suggest that the children's sexual abuse disclosures emerged late in the case investigation and were not spontaneous. The children's claims transformed over time, in terms of central details including whether they were abused, who they were abused by and where they were abused. The three children initially denied being sexually abused; however, their statements changed over time into graphic descriptions of the sexual abuse that they had suffered, and of the abuse of other children at the JIS that they claimed to

2 The last author of this paper served as an expert witness in the case, and was therefore able to provide additional information from the case records. For this purpose, she was authorised by the defence lawyer of one of the defendants to divulge information pertaining to the defence investigation that could be pertinent to the description of the JIS case.

have witnessed.

The first time the children named a perpetrator, they mentioned a couple of the school janitors. Later, they added four janitors, two security staff, three teachers, and the headmistress of the school. The three children claimed that the occurrences of abuse took place during their school schedule. The children described a pattern in which the perpetrators drugged and sodomised the children, returning them directly back to their classrooms after the occurrences of abuse. In regard to the settings in which the alleged occurrences of abuse had taken place, the children reported different sites in every interview, including students' bathrooms, teachers' bathrooms, classrooms, staff offices, and a 'secret' room that – according to the available information – the investigation team was unable to locate.

As the investigation progressed, the children's parents – who appeared to be certain that the children had been victimised – continued to attempt to extract a full disclosure from the kindergarteners. Not only were the parents certain that the children had been abused, but the police also did not explore alternative scenarios in which the accused staff were innocent. The children were sent to clinical psychologists hired by their respective families to evaluate the possibility of sexual abuse. To the best of our knowledge, these therapists employed the following methods and tools: Eye Movement Desensitization and Reprocessing (EMDR) therapy[3], imaginary play, and anatomical drawings. They also used the JIS yearbook during their therapy sessions with Child 1.

Research criticises the use of such suggestive techniques when attempting to facilitate recall, as they are likely to foster false memories (see: EMDR therapy, Houben et al., 2018; anatomical drawings, Otgaar et al., 2012). In our opinion, many of the questions that were directed to the children from different sources, including the forensic interviews the children underwent, failed to meet scientific guidelines that preserve children's integrity. The lack of scientifically sound tools and the use of techniques known to be suggestive may have reduced the children's ability to produce reliable statements. The case materials indicate that the investigation team could not find physical evidence in the pertinent sites to corroborate the allegations.

With respect to the potential offenders, the investigation team did not specify why the police dismissed the headmistress, a teacher, and a security staff member as suspects, even though the children had disclosed that these people were also involved in the abuses the children had allegedly suffered. For example, one child gave the following statement:

3 'EMDR requires that the client engage in imaginal recall of the disturbing event and focus on associated affect, negative cognitions, and body sensations while performing rapid tracking eye movements by following the repetitive motion of the therapist's hand. After the eye movement set, which usually lasts for about 20 s, the client briefly reports on any changes in the image or concurrent experiences. The client then engages in the next set of eye movements during which she or he is to focus on any new, spontaneously generated associations. This cycle of images in conjunction with eye movement followed by the client's feedback is continued until the client feels comfortable and reports that the original memory fails to elicit discomfort. At this point, a positive cognition is paired with the original scene by having the client imagine the original scene and rehearse the positive statement covertly, while simultaneously engaging in eye movement.' (Feske, 1998, p. 171)

> Sometimes the boss picked me up in class and took me to Miss Evil [headmistress] and I was taken by the boss and Miss Evil to the upstairs room the boss climbed up the stairs, they said I was bad and must be punished. Sometimes Miss Evil picked me up in the class room and then we went to the boss' room and I was taken by Miss Evil and the boss up the stairs, and then I was punished, the punishment was that the boss' birdie [word for genitals] was put into my poo hole. (Child 1, Police interview, 16 May 2014[4])



The children also describe some members of the JIS staff as heroes who interrupted the scenes in which the children were being assaulted. However, it is not clear from the materials we examined whether the police ever interviewed these staff members as witnesses. In fact, the evidence that corroborates the guilty scenario in the JIS case does not include any eyewitness statement besides the children's and their families (Neil Bantleman, aka Mr. B., 2017). This lack of corroborative testimony is striking, given that some of the children's accounts indicate that there were witnesses to the abuse. Child 1 describes being rescued by a schoolteacher (called 'Teacher A.' from now on), but we were unable to verify that this person in fact exists:

> Bad Teacher did the same thing as the boss, putting his birdie into my friends' poo hole, while Miss Evil [headmistress] was making blue soda and holding a camera and then leave. Teacher A. kicked the door open and saw me and my friends naked, then Teacher A. helped me and my friends with our clothes. Teacher A. was also angry with the boss, then went out, and came back with Miss Evil, still angry and then leave. (Child 1, Police interview 16 May 2014, Jakarta Police records gathered by the defence of the case)

In the police files of the JIS case, there is no mention of schoolteacher Teacher A.'s statement. The headmistress denied ever witnessing or being responsible for sexually abusing the children. To the best of our knowledge, the camera Child 1 mentioned was never found by the police in the warranted searches conducted at the school.

The Supreme Court of Jakarta convicted six outsourced janitors and two JIS schoolteachers of sexually abusing the three kindergartners solely based on the statements from the children, their family members, and the children's psychotherapists. One janitor died in police custody during the initial police interrogation. The other accused janitors and teachers received sentences that ranged from 7 to 11 years in prison. In 2019, after diplomatic intervention, one of the teachers, who was originally from Canada and had been sentenced to 11 years in prison, was granted clemency from the Indonesian president's office, after accepting a guilty plea. The other convicted teacher,

4 Jakarta Police written record in verbatim.

an Indonesian citizen, refused to admit guilt to be granted clemency. He remains in prison, where he has been since April 2014, as do four of the janitors. The only female janitor in the case, who was able to afford a defence lawyer, while the others were not, was released in the beginning of 2018, after spending close to 4 years in prison. Table 2 summarises the major legal events in the JIS case.

**Table 2.** Timeline of Formal Events in the JIS Case

| Time | Event |
|---|---|
| 24/03/2014 | Mother 1 reports to the police that her son was sodomised in the JIS toilet. |
| 01/04/2014 | The investigation team reconstructs the alleged crime scenes at the JIS with Mother 1 and Child 1's help. |
| 15/04/2014 | Mother 1 hosts a meeting for about 200 parents from JIS students. During the meeting, Mother 1 potentially misinforms parents about her son's medical tests for genital herpes[5]. Mother 2 claims that her child was also sexually abused. Both Mother 1 and Mother 2 encourage parents to question their children at home. |
| 16/04/2014 | Mother 2 files a report to the police stating that, in addition to the janitors, her child was abused by two of the JIS teachers. |
| 21/04/2014 | Mother 1 files a civil accusation to South Jakarta district level court, asking for $12 million USD as compensation. |
| 24/04/2014 | The Indonesian National Commission for Child Protection and The National Board for Eyewitness and Victim Protection make a press release stating that it was a fact that the children in the JIS case had been sexually abused. |
| 26/04/2014 | The police arrests six janitors who worked at the JIS. One of the janitors is found dead and covered in bruises at the police detention toilet. The police state that he committed suicide by drinking floor cleanser liquid. |
| 28/05/2014 | Mother 1 raises the initial lawsuit of $12 million USD against JIS to $125 million USD. |
| 11/06/2014 | Mother 1 files a new police report alleging that, in addition to the janitors, two teachers of the JIS also sexually abused her son. |
| 30/06/2014 | Police reconstruction of the crime scene in the JIS's toilet: In the presence of the police as well as Child 1 and Child 2, Mother 1 and Mother 2 narrate how the occasions of sexual abuse would have occurred. |
| 22/12/2014 | The South Jakarta District Court convicts the outsourced janitors accused of sexually abusing Child 1 at the JIS. Their sentences range from 7 to 8 years in prison. |
| 02/04/2015 | The South Jakarta District Court convicts teachers Bantleman and Tjiong of sexually abusing Child 1, Child 2, and Child 3 at the JIS, with sentences of 10 years in prison. |
| 25/02/2015 | The Jakarta Province Higher Court rejects the janitors' appeal. |
| 22/02/2016 | Bantleman and Tjiong's defence lawyer files a request for appeal to the Supreme Court. |
| 24/02/2016 | The Supreme Court rejects the teachers' appeal. |
| 10/08/2015 | The South Jakarta District Court decides that JIS is exempted from Mother 1's lawsuit of $125 million USD against the school. |
| 11/10/2018 | Mother 1's attorney refiles a lawsuit to the civil court against the JIS. |
| 19/06/2019 | The Office of President Joko Widodo grants Bantleman clemency. Bantleman and his wife repatriate to Canada. |
| 23/07/2019 | The South Jakarta District Court rejects the civil lawsuit of Mother 1, finding in favour of the defendants: the teachers, cleaners, the JIS, the ISS, and the Ministry of Education of the Republic of Indonesia. According to the judge in charge of the lawsuit, the lawsuit was not accepted on the grounds that the case against the school was not clear, and the case presented by the prosecution was not enough to prove any wrongdoing on the part of the school. |

5 Some of the janitors who were suspects in the case had tested positive for genital herpes. Despite Child 1's negative test for genital herpes, Mother 1 used the information on Child's 1 test that reported high levels of a certain group of immune cells that indicate body infection to argue that Child 1 had contracted genital herpes.

## How Did the Jakarta International School Case Emerge?

In early March of 2014, a large-scale child sexual abuse case was reported to the authorities at the American Nicaraguan School (*Glance at the Career of Teacher William Vahey*, 2014). The accused perpetrator, William Vahey, was an American schoolteacher with a prior record of child molestation from 1969. Beginning in 1973, Vahey taught abroad in wealthy international schools, including a stint working at the JIS from 1992 to 2002. In March 2014, Vahey was working at an international school in Nicaragua, at which time a co-worker found Vahey's thumb-drive, which contained sexual images of children. The co-worker turned the thumb-drive over to the authorities. After learning that his photographs had been discovered, Vahey committed suicide.

The Vahey case received broad international media coverage, which came to the attention of the JIS community. Because Vahey had worked as a schoolteacher at JIS, the school felt the urge to address the possibility of JIS students still being victims of abuse by the school staff (Paddock, 2014). On the 11th of March, the JIS sent parents a memo of their 'Child Protection Policy' with information on different types of child abuse. The policy presented information on different types of abuse and potential behavioural indicators of abuse and asked parents to be alert for such behaviours. In Table 3, an excerpt from the JIS protection policy is presented.

**Table 3.** Excerpt of JIS Policy on Behavioural Indicators of Sexual Abuse Sent on 11 March 2014

| **Possible Physical/Behavioural Indicators** |
|---|
| Sexual drawings or language |
| Bedwetting |
| Eating problems such as overrating or anorexia |
| Self-harm or mutilation, sometimes leading to suicide attempts |
| Saying they have secrets they cannot tell anyone about |
| Substance or drug abuse |
| Suddenly having unexplained sexual knowledge, behaviour or use of language not appropriate to age level |
| Unusual interpersonal relationship patterns |
| Venereal disease in a child of any age |
| Frequent urinary tract infection for young children both male and female |
| Evidence of physical trauma or bleeding to the oral, genital or anal areas |
| Difficulty in walking or sitting |
| Refusing to change into physical education clothes, fear of bathrooms |
| Child running away from home and not giving any specific complaint |
| Not wanting to be alone with an individual |
| Pregnancy, especially at a young age |

Shortly after the JIS sent this protection policy letter to parents, Mother 1 reported being alarmed about behaviours she saw in her 6-year-old child, who was a kindergarten student at the JIS. In the police reports and trial testimony, Mother 1 claimed that Child 1's behaviours were indicative of child sexual abuse. Mother 1 reported that Child 1 had episodes of diurnal and nocturnal enuresis and that, at times, he refused to take off his trousers to shower or use the bathroom. Furthermore, Child 1 seemed to be jealous towards his younger brother and was becoming a fussy eater, consequently losing weight. Mother 1 became suspicious that a sex ring could be operating at the JIS, with her son and other JIS students as victims. As observations as such can be misleading, in the section below, we discuss the validity of parental observations of so-called *behavioural indicators* of sexual abuse.

## Problematic Use of Child Sexual Abuse Checklists

Children sometimes show behaviours that cause alarm to their parents or pertinent professionals. For example, as a normal part of child development, children often show interest in their genitals and in the genitals of others (Poole & Wolfe, 2009). Occasionally, parents worry that a young child's interest in sexual body parts might be indicative of sexual abuse and that the child is 'acting it out' (Cromer & Goldsmith, 2010). In the case of diurnal enuresis, a young child may wet his or her pants for reasons other than an active avoidance of bathrooms (e.g. genetics, failing to remember to take a break, bad dreams, etc.; Foxman et al., 1986). Even if a child demonstrates a genuine fear of bathrooms, other factors may explain the fear. For example, children often fear loud noises such as a toilet flushing (see Stewart, 2011).

Psychotherapists in the 1980s and 1990s used child abuse checklists or behavioural indicators in efforts to identify victims of abuse who had not yet come forward to disclose their experiences (e.g. Sgroi, 1982). Such behavioural checklists are still frequently found on child protection websites (e.g. the World Health Organization). The search term 'warning signs for child sexual abuse' yields over 24 million hits (23 July 2020, Google search). However, the behaviours on these checklists are not specific to abuse. That is, children show many of the same behaviours for reasons outside of abuse. For some items on the JIS checklist (e.g. fear of bathrooms), it is questionable whether there is any relationship to abuse. Because child abuse checklists or behavioural indicators are not sound instruments, psychologists have long denounced their use to identify sexual abuse (see Franklin et al., 2018; Hibbard & Hartman, 1992).

No cluster of behaviours is seen among all or even most abused children (Finkelhor et al., 2014). Sexual abuse often happens unnoticed and is not detected in childhood unless the child decides to come forward and tell someone (McGuire & London, 2020). Most behavioural sequelae derived from child sexual abuse may also take place due to other environmental circumstances (e.g. overly strict and controlling parents,

parents' divorce, excessive television, and video-game usage), due to genetics, or through a complex interplay of environmental factors and genetics (Chaffin et al., 2002). In short, no psychological symptom profile is specific to child sexual maltreatment (see London et al., 2005, for a review). Like many other abuse checklists that are widely available online, the behaviours listed in the JIS protection policy are typical childhood behaviours (i.e. enuresis, eating problems, jealousy of a younger sibling) that would not necessarily be linked with child sexual abuse (Bernard-Bonnin, 2000).



## Allegations and Investigation Bias

Our reading of the case materials indicates that the allegations against the janitors and teachers emerged gradually after the JIS sent a copy of the Child Protection Policy to parents. Mother 1 was the first JIS parent who alleged her son had been sodomised at school based on normative childhood behaviours[6] (*Bantleman vs. Indonesia*, 2015; 2017; defence records). Despite the police reaching out to several JIS families, to our knowledge, only two other parents developed the same concerns and attitudes as Mother 1.

This aspect of the JIS case differs from other unproven cases of sexual abuse revictimisation that occurred in the 1980s and 1990s, in which *hundreds of families* seemed to become convinced that their children had been sexually abused after one parent first became suspicious (e.g. McMartin Preschool case; Campo et al., 2006). According to the police records gathered by the defence of the accused in the JIS case, at the beginning of the JIS case investigation, the three kindergarten boys repeatedly denied having suffered or witnessed any inappropriate sexual activity. However, in our reading of the materials, it appears that their families construed their denial as reluctance to disclose a traumatic experience.

Based on the Court reports that are publicly available (*Bantleman vs. Indonesia*, 2015; 2017) and the defence records, it seems that none of the children's families, therapists or police considered that the children's initial denials of being sexually abused had indeed meant that they had not been abused. Given that the police and therapists do not appear to have considered an alternative scenario in which the denials were authentic, they may have been biased towards extracting a narrative from the children the corroborated the parents' suspicions. This is not to say that the investigators were acting in bad faith; rather, they may have been disposed to seek confirmatory evidence.

The children's eventual allegations did not originate in isolation from one another. Rather, in our opinion, the claims and events that compose the JIS case illustrate a prototypical example of 'snowballing' and 'cascade bias' (Dror et al., 2017). Dror et al. (2017) explain how human cognitive biases can distort decision-making in the forensic context. Bias cascade effects are defined as irrelevant information cascading from one stage of an investigation to another. Also, in bias snowball effect situations, '…bias

6 E.g., being jealous of his younger brother and being a picky eater

increases as irrelevant information from a variety of sources is integrated and influences each other' (Dror et al., 2017, pp. 832–833).

From a speculative perspective, Mother 1's belief about her son's developmental behaviours may have affected how the police officers interacted with the suspects and the alleged victims. In addition, the defence records show that Mother 1 shared her concerns with Mother 2 and Mother 3, who eventually obtained similar information from Child 2 and Child 3, respectively. According to the court records (*Bantleman vs. Indonesia*, 2015; 2017) and the defence records, Mother 1, Mother 2, and Mother 3 took their children to the same therapists. Reportedly, the three children attended various therapy sessions together. In these sessions, the children were not prevented from hearing their respective mothers explaining to the therapists what the children had supposedly experienced.

With the build-up from therapy appointments, the snowball effect may have burgeoned. That is, the children's initial denial of sexual abuse became graphic descriptions of how several men and one woman, one after the other, had sodomised them and other JIS students on multiple occasions. One of the therapists who had treated the children disclosed the following: 'Other than himself, Child 3 admitted that he once saw three other friends, plus Child 1 and Child 2, in Neal Bantleman's room on the upper floor, after Child 3 had been sexually abused' (Psychotherapist dossier, 21 July 2014; *Bantleman vs. Indonesia,* 2015; 2017).

Child 2's first denials were as follows: 'No, he [Mr. Bantleman] did not punch me, he also did not insert his bird' (Police records, 18 May 2014). About a month later, Child 2 made his first (registered) disclosure of child maltreatment to the Indonesian Police:

> At that time, I was abused and harmed by bad people, then I report it to my homeroom teacher, then I was escorted to the room of the principal, then I was escorted to the upper room, then in that room I was taken to a room where there was Mr. Bantleman. In that room, the principal accompanied me. That room has a window that always closed by curtain. I cannot see outside. I sat on a green couch and report to Mr. Bantleman that I was harmed by Agun (one of the janitors) and Mr. Tjiong, but Mr. Bantleman said that I am liar and bad boy. Then I was hit in my stomach, and I was pushed to the floor by Mr. Bantleman and he opened my pants. Then Mr. Bantleman put his pity-pity to my anus, and it felt hurt very much. After that I wore my pants by myself, then Mr. Bantleman said, 'if you dare to tell to other people, I will kill your mom,' and he told me to go out. (Police records, 26 June 2014)

Furthermore, a particularly curious element emerged in the children's statements:

Police: When Child 1 refused to drink, did Miss Evil (headmistress) or the boss (schoolteacher Bantleman) get angry with Child 1? Explain.
Child 1: They were not angry with me, but the boss took a magic stone that the boss took from the sky.
Police: How did the boss take the magic stone from the sky? Explain.
Child 1: The boss went like this (Child 1 demonstrated by raising his right hand and flicking his finger).
Police: After the magic stone was taken from the sky, what did the boss do? Explain.
Child 1: The boss put the magic stone into my poo hole, he said so that it wouldn't hurt if I get punished. (Police records, 16 May 2014)



Given how unrealistic the details of this account are, the police probably did not take the magical stone narrative as factual. However, it is unclear why the police did not attempt to reconcile this prominent detail with reality. For example, the magic stone could be a detail of an authentic event that was misunderstood by the child (e.g. a rectally administered drug), or it could simply be a fabrication. However, it appears that no possible explanations for this unrealistic element in the narrative were explored. In the interview from which we extracted the excerpt above (part of the records of the defence of the case), the police officer in charge did not ask follow-up questions to clarify details about the magical stone. Thus, in our view, the investigation seems to be insufficient – that is, in terms of both considering alternative scenarios and assuming that the allegations were basically true – for obtaining a clear and consistent understanding of how the abuse happened.

In our view, the therapists' opinions that were reported to the police regarding the children's behaviours and drawings often did not seem to be neutral evaluations and frequently involved substantial speculation. One of the therapists who had seen all three children asserted that a drawing of 'rain' made by Child 1 in his first session meant that he had been sexually abused. At the teacher's trial, she reported that drawings of rain indicate catastrophic events that, in Child 1's case, meant being sexually abused (*Bantleman vs. Indonesia*, 2015; 2017*)*. To the best of our knowledge, there is no scientific evidence that supports this interpretation of drawings of rain. The same therapist used anatomic drawing techniques with the children to help them disclose the alleged occasions of abuse. In the therapist's report, she mentioned the following:

Therapist: … I can state that the drawings he [Child 2] made are caused by sexual assault/violation.
Police: How could you know that the drawings which are shown by the

examiner indicate sexual assault/violation experienced by Child 2?

Therapist: It is in accordance with knowledge in the field of psychology. I can see it from the position of the drawing, the pressure of the pencil/pen/crayon, the size of the drawing, comparison of body and drawing specification, like eye, ear, nose, etc. (Police report, 2 May 2014)

In our opinion, no validated instruments nor psychological research support this psychologist's claims (see Otgaar et al., 2012). Consistent with the snowball and cascade bias effect, the bias of the professionals who were partially in charge of the JIS investigation may have impeded the children from providing reliable accounts and consequently hindered the authorities from considering scenarios in which the accused were innocent.

## Suggestibility and False Reporting as a Result of Poor Interviewing Practice

Based on the results of decades of research on child suggestibility, scholars have developed specialised interviewing protocols to account for children's susceptibility to fall prey to external suggestion (Bruck & Ceci, 1999; Ceci & Huffman, 1997; Endres, 1997; King & Yuille, 1987). Here, *suggestibility* refers to the 'degree to which children's encoding, storage, retrieval, and reporting of events can be influenced by a range of social and psychological factors' (Ceci & Bruck, 1993. p. 404). Research has shown that young children are inclined to please adults' expectations (Bruck & Ceci, 1999), and that children may confabulate stories they did not experience so that they agree with their peers (Principe, Cherson, DiPuppo, Schindewolf, 2012). Furthermore, while exposing a child to repeated interviewing can assist their recall and prevent them from forgetting *true* events (Howe & Courage, 1993), children who are repeatedly questioned about a false (i.e. non-experienced) event are prone to memory distortion (Bruck et al., 2002; Leichtman & Ceci, 1995).

Because the initial reports of child sexual abuse in the JIS case were not a product of the children's spontaneous accounts, the number of interactions the children had with adults may have been problematic. According to the reports available to the authors (both the public court report and the defence report), during their interactions with the children, the adults sought to collect from them a confirmation that the children had been sexually abused. When children are induced to report maltreatment by a biased interviewer, the reliability of the report can be compromised (London et al., 2005).

The police file of Child 1 states that he was questioned by the police at least 13 times, had six invasive medical examinations for child sexual abuse, and attended 52 psychotherapy sessions. In Child 2's case, he was questioned by the police at least five times and had six invasive medical examinations for child sexual abuse. Child 2 had 16 documented psychotherapy sessions. Child 3 was questioned by the police once and

had 10 documented psychotherapy sessions. We believe that these numbers indicate the minimum sum of the children's interactions with adults and with each other; they also illustrate what the children went through in the JIS case. Furthermore, Mother 1, Mother 2, and Mother 3 often stated to the police, the therapists, and other JIS parents that they questioned their children frequently at home. The mothers also stated that their families visited each other and attended therapy sessions together, and that the children often interacted with each other and played together (*Bantleman vs. Indonesia,* 2015; 2017). In the context of a legal case, it can be problematic to allow allegedly victimised children to exchange information with one another due to children's potential propensity to misattribute the source of information, consequently contaminating each other's reports (Candel et al., 2007).



Children can be highly influenced by natural conversation with their peers and parents (Principe & Schindewolf, 2012). In the JIS case, the children's mothers, therapists, and other family members interviewed them repeatedly over the course of the investigation. Principe and colleagues (2013) conducted a study showing that, when misinformed, mothers induced their children to report memories of a false event. Of the children in that study, 60% were affected by the misinformation acquired by their mothers and their suggestive interviewing manner (for a review, see Lawson et al., 2018).

The JIS children were interviewed together, played together and even attended therapy sessions together. Thus, a combination of peer exposure and suggestive interviewing may have led the children to falsely claim to have witnessed activities that only their peers had (or were believed to have had) observed. Research shows that such interactions may be problematic. After repeatedly hearing from their peers what they had experienced, children who did not experience the same events have been found to report details and information comparable to what their peers had reported (Principe & Ceci, 2002).

As previously discussed, children's propensity to suggestibility combined with misleading information may result in false reports and potentially false memories (Ceci & Bruck, 1995; Ceci & Friedman, 2000). Specific to the children in the JIS case, research shows that 6-year-olds – compared with 3-year-olds – are more likely to draw and falsely report inferences of obscure situations, often mistaking them for actual events (Principe et al., 2008). In our opinion, it is plausible that rumours of sexual abuse in natural conversations with their parents and peers could have become a source of false memory for Child 1, Child 2, and Child 3. Moreover, research shows that repeated exposure to suggestion may also lead to false memory formation (Zaragoza & Mitchell, 1996).

When the situation is held against the literature on this topic, it becomes clear that the JIS children were exposed to a variety of elements that are known to elicit various types of false memories (as well as false reports in the absence of false memories). False memories can be formed by active external influence through misinformation about the

details of real events (Loftus, 2005). Furthermore, false memories of entirely false events, driven by suggestion and social influence, can be implanted into the memory (Loftus & Pickrell, 1995). False memory implantation refers to a specific situation in which a person remembers an entirely false autobiographical event after one or more rounds of suggestive interviewing.

In general, false memories can be formed for both plausible and implausible events. Principe and Smith (2007) demonstrated that children's belief of an unreal phenomenon (i.e. the tooth fairy) can engender false memories that are in line with their imagination. Otgaar and colleagues (2009) convinced 7- to 8-year-old and 11- to 12-year-old children that an unidentified flying object (UFO) had abducted them when they were younger. Another implausible false narrative that was used in a different study with undergraduates suggested to the participants that they had witnessed a person being demonically possessed (Mazzoni et al., 2001). In another study, Strange and colleagues (2006) had 6- and 10-year-olds remember having had a cup of tea with Prince Charles.

Based on the materials available to the authors, there is an indication that, in the JIS case, the parents and the police repeatedly presented to the three children the narrative of being sexually abused as something that was highly likely to have happened to them. According to the available information, none of the parents and professionals allowed the children to offer a different scenario, even when the children denied being abused. In our view, one possibility in the JIS case is that the children falsely reported being sexually abused. It is possible that the children in this case formed rich false autobiographical memories of being sexually abused after being overly exposed to strong social influence (see Zaragoza et al., 2011). There is no direct evidence of false memory formation in this specific case. However, the literature on false memory implantation demonstrates that many of the circumstances the JIS children were exposed to are known to elicit detailed yet inaccurate reports.

Furthermore, parents are generally inadequate interviewers when it comes to collecting reliable accounts from their children, as they are likely to prompt children to disclose non-factual information (Bruck & Ceci, 1999; Ceci & Huffman, 1997; Korkman et al., 2014). We believe that the alleged victims in the JIS case seem to have been affected by a series of different influences, such as parental interviewing, misleading forensic interviewing, rumourmongering, social pressure, and other types of suggestion. Below are some excerpts from the registers of the police interview that were made available to the defence, which illustrate how the parents in this case interviewed their children. Moreover, the excerpts below imply that Mother 2 and one of the therapists were resistant to believing in the JIS children's denials:

Mother 2: Then until the incident happened to Child 1, Mother 1 spoke to the press on 14 April 2014 in the morning, she held a press conference; then I read, in that event, he [the reporter] said, how did she [Mother

> 1] know that her child is changed, [he] wet his pants, [he was] afraid to go to the toilet, and [he] said there is a bad guy that does something bad to him. Then I said, 'this is similar to what happen with my son, and every time I asked him, he answered differently, [he] never give me a consistent answer. (Mother 2's testimony, 23 December 2014)

Mother 2 tries to convince Child 2 to speak about the alleged sexual abuses that Child 2 had presumably suffered in school; however, Child 2 does not confirm that he was sexually abused. Instead, Child 2 says that he was physically punished at the JIS. Mother 2 reported the following dialogue between herself and Child 2:

> Child 2 said, 'How did you know that they are cleaners?', then I said, 'I did not say that they are cleaners, I only said the man who wears a blue shirt, is that cleaner?' Child 2 said, 'Yes, mom, they are cleaners', and I said, 'What did they do to you?'
>
> He [Child 2] is beaten, tortured, and I asked, 'Is there anything else?' He did not want to say it. Then I said, 'Okay, Child 1 your friend, has been assaulted like this, he said that there is a bad guy who put his 'pity-pity' into his butt'. 'Eww! So gross, mom, about the thing that you tell me that' Child 2 said. Then I told him, 'I only want to know why you are afraid to go to the toilet; you don't want me to clean you up, you don't want to go to the toilet because there is a bad guy, you just told me that you are afraid; therefore, I just want to know', Child 2 said, 'No, mom, no, I am kicked and beaten, did you remember that there is a bruise on my leg, and mom asked me several times and I said that I fell [unclear]'. (Mother 2 testimony, 23 December 2014)

Therapist: At the third meeting, Child 2 initially did not want to talk, then Child 2 finally wanted to. In fact, for the first time he said that he was confronted with 3 people, but Child 2 did not admit that he experienced sexual violation. (Therapist police statement, 5 May 2014)

According to the research discussed, factors present in the JIS case, such as parent interviewing, misleading forensic interviewing, rumourmongering, and social pressure, have the potential to create false reports or even elicit false memories.

## Line-up and Face Recognition

When questioning Child 1 at home, Mother 1 showed Child 1 photos of the 30 different JIS janitors (*Bantleman vs. Indonesia*, 2015; 2017*)*, employing a line-up and face recognition procedure. Despite lacking training for the task, Mother 1's aim was to have Child 1 point out which of the 30 janitors had sexually abused him. When making an identification through a line-up in which the perpetrator is present, children can be as accurate as adults (Harvard, 2014). However, when presented with a line-up in which the perpetrator is absent, children tend to make false identifications, since they often align with social demand and want to give a positive response (Harvard, 2014).

Mother 1 shared yearbook photos with Mothers 2 and Mother 3, who also used the photos to question Child 2 and Child 3. Mother 2 reported showing her son the photos of the people Child 1 had identified as the perpetrators. On July 5, Child 3 attended a line-up at the police station, with 34 photos of janitors at the school that included the suspects pointed out by Child 1 and Child 2.

On the same day, Child 3 underwent a 'condom line-up' with the police, in which he was shown differently coloured condoms and asked to select the colour worn by his alleged perpetrators[7]. This procedure presupposed that Child 3's supposed abuser wore a condom; however, the investigation did not base the line-up on information Child 3 had divulged. In fact, Child 3 denied being abused and consequently having seen colourful condoms. Hence, it could be concluded that Child 3 chose a coloured condom because his only option was to choose a condom from the line-up.

On 21 July 2014, Child 1 and Child 2 went through an 'indirect confrontation' with Mr. Tjiong and Mr. Bantleman. The police placed both perpetrators in a live line-up with two other men. The documents do not specify which line-up happened first, Mr. Bantleman's or Mr. Tjiong's. Both Child 1 and Child 2 'recognised' Mr. Bantleman and Mr. Tjiong. Child 1 and Child 2 had been shown the photos of both teachers multiple times prior to the line-up; hence, it might be argued that the evidence gathered in this procedure was potentially unreliable.

## Contemporary Best-Practice Standards in Forensic Interviews with Children

If behavioural indicators are not reliable markers of child sexual abuse, how can professionals evaluate cases in which abuse concerns arise? Expert witnesses and practitioners can assess the evolution of the statements. That is, they can determine how much the children's statements have changed over time in comparison with their original statements (London et al., 2005). To examine the reliability of a child's report, it is critical to determine whether the child's first statements were made relatively

7 Previous to this line-up, Father 3 reported having shown Child 3 a condom at home to attest whether he was able to recognise it. The police reports do not state whether Child 3 did or not recognise the condom his father showed him at home.

spontaneously to neutral interviewers. If children's accounts of sexual abuse are induced rather than spontaneous, they are less likely to be factual (London et al., 2005). Children's statements are often unreliable when elicited by parents or interviewers who hold pre-existing beliefs about what happened. Such *a priori* beliefs can introduce a range of suggestive influences into informal conversation or formal interviews (e.g. leading questions).

However, when interviewed under appropriate circumstances, children can give accurate reports of prior events. Lamb and colleagues (2007) developed the evidence-based NICHD protocol and provided evidence to support its usefulness (e.g. see Lamb & Brown, 2006; Lamb et al., 2007; Roberts et al., 2004). Another protocol, the Cognitive Interview (Holliday, 2003), was also scientifically derived and has been extensively used by various agencies to interview children in different countries. Moreover, research shows that, if children have been influenced to disclose false information, even if the children's subsequent statements are collected by a neutral interviewer, the original false statement is likely to continue, which shows the influence of the earlier interviews (for reviews, see Bruck & Ceci, 2004; Bruck & Melnyk, 2004; Ceci & Bruck, 1993). There are various conditions and types of questions that can undermine children's memory and accuracy (see Lavoie et al., 2021).

To investigate the possible influence that some of these techniques could have on children's suggestibility, scientists have studied interview techniques derived from those used by psychotherapists (focused on the recovery of repressed memories) and law enforcement (e.g. London & Kulkofsky, 2010; Pipe et al., 2004; Poole & Lamb, 1998; also see Ackil & Zaragoza, 1998; Finnilä et al., 2003; Garven et al., 1998; Lindberg et al., 2000; Orbach et al., 2000; Sumampouw et al., 2019; Warren & McGough., 1996; Zaragoza et al., 2001). Based on this large corpus of research, experts in the field of children's memory, adult memory, and interview techniques generally agree on the following four general principles:

(1) Careful attention must be given to the first and earliest interviews of children, before any suggestive interviewing has occurred, and before there is a lapse of time that would result in normal memory decay related to the allegation. Informal conversations that took place with the child before formal interviews must also be considered (Korkman et al., 2014; Lamb & Brown, 2006; London et al., 2005; London & Kulkofsky, 2010).
(2) Suggestive interviews are not characterised solely by leading questions; rather, a central feature of suggestive interviews is a biased interviewer who uses a variety of explicit or implicit suggestive techniques to elicit answers that are consistent with their primary hypothesis (Poole & Lamb, 1998; Roberts et al., 2004; Dror et al., 2017).
(3) Once children have been exposed to biased and suggestive interview techniques,

in the absence of clear corroborating evidence, it is difficult to distinguish between true and false reports (Bruck et al., 1997; Leach et al., 2004; Leichtman & Ceci, 1995; Pezdek et al., 2004).

(4) Children should be formally interviewed as soon as possible after suspicions arise, and the interviews should be video recorded. Video recordings allow the factfinders to assess the degree to which a child's response is a product of either neutral or suggestive interviewing practices (Holliday, 2003; Lamb et al., 2007).

To summarise, one way to assess the reliability of a child's report is to determine whether the child's first statement was spontaneously made to neutral interviewers. Furthermore, if a child's first statement was elicited by the child's parents or interviewers who held pre-existing beliefs about the occurrence of a particular event, the child's statement may not be reliable. This is particularly the case when the interviewer has used a range of suggestive interviewing techniques, such as the use of an anatomic doll or closed questions. Without clear evidence, it is never possible to say whether a statement is true or false; however, it is possible – as we do here – to evaluate whether the circumstances of the interviews included factors that are known to improve or damage the reliability of statements. In the present case, we have evaluated the processes by which allegations were made by the complainants from the time at which any concern about inappropriate sexual activity was raised to the time of the trial. In our view, the interviews conducted with the JIS children were not in line with best practices derived from the scientific literature.

## Investigation of Alternative Scenarios

Mother 1's initial concerns might have created a cascade bias effect (Nakhaeizadeh et al., 2017) that trickled down to affect the subsequent practices and interpretations given by the medical, psychological, and legal professionals working in the prosecution of the case. If this was the case, it is quite unlikely that different resolutions to the case could have emerged from such partial investigation. In the absence of clear medical indications of sexual abuse or positive test results for sexually transmitted diseases, there was no solid incriminating evidence against the JIS staff. The investigation team could not find supporting evidence, such as witnesses who saw the children being sexually abused.

Bias among expert witnesses and police officers can potentially be mitigated by working with alternative scenarios (see Otgaar et al., 2016). In cases concerning the reliability of testimony, they can work with at least two different scenarios (i.e. guilty and innocent), offering different theories for a given crime. The 'guilty' scenario considers the victim's allegation to be authentic, while the 'innocent' scenario assumes that the victim's statement is not authentic, but rather a product of errors (e.g. false memories or deception). When using at least two scenarios (i.e. guilty and innocent), evidence (or

the lack of it) gathered from different sources can be considered in relation to the two scenarios. This way, information is less likely to be dismissed because it does not match the scenario the investigation believes to have taken place. As the investigation proceeds and more evidence corroborates one of the scenarios, the investigation team has more clarity on which scenario is most likely to have taken place. Such an approach might protect practitioners of the legal system from different types of biases during their work (e.g. confirmation bias; see van Koppen & Mackor, 2019).



## Summary of Recommendations and Final Remarks

Child abuse is a major public health problem worldwide (World Health Organisation, 2020). According to decades of laboratory and field research, reliable forensic statements from children emerge under certain conditions: to the extent that children come forward relatively spontaneously and make statements in their own words (without a motive to lie), the reports are likely to be reliable (see Hritz et al., 2015). Children can remember and provide narratives about experienced events, even after weeks of delay and even involving traumatic events. In our opinion, the JIS case does not contain these markers of reliable reports. To the contrary, in our view, the JIS case features many factors that are known to have the potential to produce unreliable statements.

When adults hold beliefs that certain events have occurred, they are likely to tailor their questions to children in a manner that may elicit statements that align with the adults' prior beliefs (e.g. Bruck et al., 1999; Principe et al., 2013). Children may assent to the adults' suggestions while initially knowing that the misinformation is inaccurate; however, over time, the children may come to embellish the narrative of the suggested events, accepting the events or even adopting the events into their autobiographical memory (Poole & Lindsay, 2001; Poole & White, 1991). Under such circumstances, children's statements are rendered unreliable (London et al., 2010). That does not mean that coercive interviews are incapable of eliciting true statements. Rather, such techniques do not yield diagnostic information disclosures; it is not possible to distinguish between accurate and inaccurate information gained from such techniques. When interviewed later by neutral interviewers, children often continue to report the false events (Hritz, 2014). That is, once damage is done to the reliability of a child's testimony, that damage is often irreparable.

Witnesses' statements are crucial in legal investigations, in the absence of other leads (Albright, 2017). Myriad psychological processes play pivotal roles in the development of such statements. Practitioners of the legal system could benefit from education in the science of memory, especially in regard to knowledge about the different types of false memory formation and how they are elicited (see Ost et al., 2013). In many countries, not just Indonesia, the practices used in child interviewing are not in line with well-established scientifically derived interviewing principles (see Otgaar et al., 2019).

When children produce false reports based on suggestive interviewing techniques,

these reports may well seem to be superficially credible; however, there is no reliable method to distinguish between true and false statements under such circumstances (Ceci et al., 2007). Furthermore, when children's testimony is the only evidence – which is often what occurs in child sexual abuse cases – police officers may benefit greatly from employing alternative scenarios within a case investigation in order to avoid wrongful convictions. Adequate training is a tangible strategy to help minimise the miscarriage of justice in legal cases. The legal system should ensure that, when people are accused of crimes, the investigation is carried out in a way that reliably verifies that the crime occurred and that the accused are indeed the offenders.

# Chapter III

## Are Children Better Witnesses than Adolescents? Developmental Trends in Different False Memory Paradigms

Calado, B., Otgaar, H., & Muris, P.

*This chapter corresponds to the following paper:*
Are children better witnesses than adolescents?
Developmental trends in different false memory paradigms.
Journal of Child Custody, 15, 330–348, DOI: 10.1080/15379418.2019.1568948

**Abstract**

This study compares older children's (11- and 12-year-olds) and adolescents' (14- and 15-year-olds) vulnerability to false memory creation using two different methods (i.e. the DRM and memory conformity paradigms) involving neutral and negative stimuli. In line with previous research, a developmental reversal effect was found for the DRM paradigm: when employing this method, children displayed lower false memory levels than adolescents. However, when using the memory conformity paradigm, the opposite pattern was found, with adolescents forming fewer false memories than children. These findings indicate that, in a co-witness context, adolescents are less prone to memory errors than children. The emotional valence of the stimuli used in both paradigms did not notably affect the production of false memories. There was no statistically significant correlation between false memories as measured by the DRM and the memory conformity paradigms. Altogether, the current study indicates that there is no single type of false memory, as different experimental paradigms evoke different types of erroneous recollections. In addition, our study corroborates past findings in the literature concerning the issue of developmental reversal, strengthening the idea that, under certain circumstances, children might indeed be better witnesses than adolescents.

***Keywords:*** False memory; developmental reversal; memory conformity paradigm; Deese-Roediger-McDermott paradigm; children and adolescents.

## Opening

Eyewitness testimonies are commonly crucial pieces of evidence in police investigations. Especially when physical evidence (e.g. DNA traces) is lacking, eyewitnesses' statements can be the determinant for triers of fact to reconstruct what happened during a crime and provide a verdict. Occasionally, a single testimony is responsible for orienting the whole course of events in a case investigation, as well as its resolution (Albright, 2017). When it comes to testifying, children are usually seen as poor eyewitnesses and are considered to be incapable of producing reliable statements. This viewpoint is based on children's heightened propensity to succumb to suggestive questions that could potentially lead to memory errors, such as false memories, which may jeopardise legal investigations (Brainerd, 2013; Bruck & Ceci, 1999; Ceci & Bruck, 1993).

However, in recent years, an accumulating corpus of research has emerged showing that children are less prone to false memories than it was previously thought, and that they even produce fewer false memories than adults under certain conditions (Otgaar et al., 2018). In the psychological literature, this finding has been referred to as the developmental reversal effect (Brainerd et al., 2008). It is notable that developmental reversal has often been documented as an effect that occurs when the experimental procedure known as the Deese-Roediger-McDermott (DRM) paradigm is used (Deese, 1959; Roediger & McDermott, 1995). The goal of the current study is to investigate whether children's and adolescents' propensity to create false memories depends on the method that is used to assess false memory. This topic is relevant when interpreting children's or adolescents' testimonies in legal cases, especially when such testimonies are the only source of available information.



## Legal Cases Where Child Suggestibility Played a Crucial Role

Research in the area of children's false memory and suggestibility started – among other factors – from legal cases in which children's accounts of sexual abuse were contested. The cases of the McMartin Preschool (Garven et al., 1998) and the Wee Care Nursery (Howe & Knott, 2015) provide good illustrations of children's suggestibility and the catastrophic consequences of subsequent false allegations. As described in Chapter I, the McMartin Preschool case occurred in the 1980s and lasted for about 7 years. Starting with one child's statement incriminating teachers for sexual abuse, the number of accusations grew to hundreds – not only against the McMartin Preschool staff, but also against teachers of other schools in the same community. After several years of investigation and deleterious interviewing of all the children involved, there appeared to be no convincing evidence against the defendants, and the case was settled without convictions.

The Wee Care Nursery case showed remarkable similarities but ended with the

accused teacher being convicted of all the sexual abuse allegations made by the young children (Brainerd & Reyna, 2005). Although these famous day care cases of alleged sexual abuse happened decades ago, comparable cases occasionally appear in more recent times (Otgaar et al., 2017). An example is the Jakarta International School case, which occurred in 2014 described in Chapter II. Teachers and cleaning staff were sentenced to imprisonment based solely on children's allegations that they had been sexually abused (Millar, 2018), although a post-hoc reconstruction of the case indicates that the claims were likely the result of suggestive questioning.

## Child Suggestibility in the Legal Context

Cases such as these have prompted researchers to seek suitable approaches for interviewing children, and to investigate what memory errors children are likely to commit (e.g. Hyman et al., 1995; Lindsay et al., 2004; Loftus & Pickrell, 1995; Loftus, 1997, 2003, 2004). An important line of inquiry that was initiated by Ceci and Bruck (1993) considers children's proneness to suggestibility, which can be broadly defined as the extent to which children's memory (encoding, storage and retrieval) can be impacted by a host of different factors (e.g. social or psychological). Early studies on children's suggestibility indeed showed that younger children are more suggestible than older children and adults. For example, King and Yuille (1987) sought to examine differences in the propensity to suggestibility among children in four age groups (i.e. 6-, 9-, 11- and 16-year-olds).

The participants were placed in a room and witnessed a man watering some plants. Afterwards, the researchers interviewed the children about their recollections of this 'live' event, by asking a number of leading questions (e.g. 'On which arm did the man wear his watch?' – although no watch was present) and non-leading questions. The 6-year-olds recalled fewer details and were significantly more suggestible than the older children and adolescents. Overall, the 11- and 16-year-olds performed similarly in terms of recollection. In addition, despite the finding that individuals are less suggestible as age increases, no statistically significant differences were found between 11- and 16-year-olds.

More recent studies on suggestibility have confirmed that younger children are more suggestible than older children and adults. For example, Sutherland and Hayne (2001) investigated age-related differences in memory retention using a misinformation paradigm. In one of their experiments, participants in three age groups (i.e. 5- to 6-year-olds, 11- to 12-year-olds and adults) viewed a video and were interviewed about it at three different time points: immediately after watching the video, and at follow-ups after 1 and 6 days. Neutral, leading, and misleading post-event information was given to the participants during the follow-up sessions. At both follow-ups (after 1 or 6 days), the children (i.e. the 5- to 6-year-olds and the 11- to 12-year-olds) incorporated more misleading information into their accounts of the video in comparison with adults. Taken

together, this research demonstrates that suggestibility and, in its wake, the occurrence of false memories follow an age-related decrease during the course of development.

## DRM Paradigm and Developmental Reversal Effect

Recent studies focusing on *spontaneous* false memories – that is, faulty recollections that occur without any suggestive pressure – have revealed a reverse developmental trend. This research typically relies on the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995), during which participants receive word lists containing associatively related words (e.g. *coffee*, *liquid*, *soda*, *swallow*, *tea*, and *water*) that converge to a non-presented theme word called the 'critical lure' (i.e. *drink*). After the participants have memorised these words, they are engaged in a memory (recall/recognition) task.

The canonical finding is that many participants falsely recall and/or recognise the critical lure as being part of the word list, which can be regarded as a false memory. Studies employing the DRM paradigm in various age groups (i.e. children, adolescents, and adults) have shown that younger participants are less likely to remember the critical lure, and thus have a lower propensity for this type of false memory (Brainerd, 2013; Brainerd et al., 2008; Brainerd et al., 2002; Goswick et al., 2013; Howe et al., 2004; McGuire et al., 2015). This has been labelled as the developmental reversal effect (Brainerd et al., 2008; Otgaar et al., 2018) and can be explained by both the fuzzy trace theory (Brainerd et al., 2008) and the associative activation theory (Howe et al., 2009; Otgaar et al., 2016). These theories predict that spontaneous false memories ought to increase with age, thus making children less prone to form false memories than adults (Otgaar et al., 2013; Otgaar et al., 2018).

More specifically, fuzzy trace theory posits that memories are encoded in two different, independent traces: verbatim and gist (Brainerd et al., 2008). As described in Chapter I, verbatim traces refer to detailed characteristics of an experienced event (e.g. each ingredient used in a recipe), which tend to fade easily with the passage of time. Gist traces relate to the underlying, general meaning of the experience (e.g. cooking) and remain available even after long periods of time. False memories occur when a person cannot retrieve the verbatim traces of an experienced event anymore, and hence relies on the underlying meaning of the experience: the gist (Brainerd et al., 2008). Since the ability to extract the gist of experiences improves with age, fuzzy trace theory predicts that younger children are less likely than older children or adults to produce spontaneous false memories during a DRM task (Brainerd & Reyna, 2002).

Alternatively, associative activation theory postulates that false memories occur due to spreading activation in an individual's knowledge base (Howe et al., 2009). When activation – concerning a specific event – spreads through a person's knowledge base, content-related details from other experiences can also be activated and erroneously associated with that event, resulting in the creation of a false memory. Associative

activation theory posits that the associations between the details formed in a person's knowledge base become stronger and more automatic with age, resulting in older individuals having a greater susceptibility for false memories.

## Memory Performance in the Memory Conformity Paradigm

Overall, there are different lines of research with regard to children's eyewitness memory, and these have produced diverging results. On the one hand, there are studies showing that children are highly suggestible and more prone to false memories (Sutherland & Hayne, 2001; for reviews, see Bruck & Ceci, 1999). On the other hand, the research on spontaneous false memories has indicated that children are less susceptible to false memories (Anastasi & Rhodes, 2008; Metzger et al., 2008). This divergence in findings implies that the concept of false memory is not unitary, and that different methods used to induce false memories produce different outcomes, which occur via different mechanisms (Bernstein et al., 2018).

To further examine this divergence, researchers should follow two important steps in their investigations. The first step consists of comparing the results of different false memory paradigms in the participants of various age groups. If developmental trends differ between paradigms, it implies that different mechanisms underlie the production of false memories. The second step is to investigate the correlations between false memories arising from different false memory paradigms. If weak correlations are found, they would confirm the notion that there are indeed different types of false memories.

Despite the lack of research properly addressing these issues, some studies have been conducted that follow these two steps (i.e. Bernstein et al., 2018; Ost et al., 2013; Otgaar & Candel, 2011; Patihis et al. 2013; Zhu et al., 2010). For example, Otgaar and Candel (2011) explored developmental trends using different types of false memory paradigms. A total of 100 young participants in four age groups (i.e. 5- to 6-year-olds, 7- to 8-year-olds, 9- to 10-year-olds and 11- to 12-year-olds) were presented with DRM word lists and a suggestibility measure called the Bonn Test of Statement Suggestibility (BTSS; Endres, 1997). In line with expectations, the results showed that the children's vulnerability to DRM false memory increased significantly with age, while false memories due to the suggestive information from the BTSS test significantly decreased with age. This finding confirms the notion that different false memory paradigms appear to tap into different types of false memories. An important finding was that Otgaar and Candel (2011) did not find a statistically significant correlation between children's likelihood of having of false memories as assessed with the DRM paradigm and the BTSS.

Another example is the experimental investigation by McGuire and colleagues (2015) in which adolescents' and adults' false memories levels were investigated using the DRM and the memory conformity paradigm (see also Gabbert et al., 2003; Gabbert et al., 2004; Wright et al., 2000). The general outline of the memory conformity paradigm

procedure is as follows: in the first stage, the participants experience a real-life event (e.g. witnessing a discussion between two people) or view a video (e.g. presentations of a robbery event). Subsequently, they meet with a confederate who has supposedly experienced the same event as the participants. Unbeknownst to the participants, the confederate has actually *not* experienced the same event and hence introduces suggestive misinformation about the event during the dyad-recall conversation with the participants.

Typically, the results show that the participants often accept the confederate's misinformation, thereby forming false memories in such a co-witness context (Wright & Schwartz, 2010). McGuire and colleagues (2015) demonstrated a developmental reversal effect for the DRM paradigm in their participants (when comparing 11- and 21-year-olds). However, they did not find the same effect for the memory conformity paradigm; in other words, children formed more false memories than adolescents and adults. Furthermore, no statistically significant relationship was found between spontaneous and suggestion-induced false memories; this again reveals that each of the methods being used to evoke false memories functions through different means.



## The Current Study

Given the paucity of research on comparing developmental trends in the false memory rates obtained with different paradigms, and the few investigations that have been carried out on the relationship between these different types of false memories, the present study further examined this topic. Following McGuire and colleagues' (2015) study, we employed the DRM and the memory conformity paradigms to investigate participants' susceptibility to producing false memories. We predicted that the DRM paradigm would evince a developmental reversal effect (i.e. children would be less susceptible to false memory creation than adolescents), while the memory conformity paradigm would result in the opposite pattern, with adolescents being less susceptible to false memories than children.

McGuire et al. (2015) tested adolescents and young adults, we conducted this study with older children and adolescents. On a cognitive level, older children have (almost) the same abilities as adolescents, which might obscure age-related differences in false memory production (e.g. McGuire et al., 2015). Nevertheless, there is still a lack of research using adolescents as witnesses (but see also Sauerland et al., 2019); therefore, we found it imperative to include this age group in a study such as this one in order to further improve our understanding of differences in memory performance throughout different development stages.

For exploratory purposes, we provided the children and adolescents with neutral and negative stimuli in both paradigms (i.e. DRM and memory conformity). It can be argued that negative stimuli are more ecologically valid, as they are more clearly linked

to the events children encounter in legal cases. Moreover, research shows that negative stimuli usually lead to higher false memory rates than neutral stimuli (Brainerd et al., 2010; Brainerd et al., 2016; Howe et al., 2010).

## Method

### *Participants*

In order to determine our sample size, we ran an *a priori* power analysis with a medium effect size ($f$ = 0.375) and high power (0.95). This resulted in a required number of 26 subjects per age group, for a total of 52 participants. In actuality, we were able to recruit 58 child participants. In the group of 11-/12-year-olds, 20 participants from a Dutch primary school were tested (mean age = 11.30 years, $SD$ = 0.47, 8 males and 12 females). In the group of 14-/15-year-olds, we recruited 38 participants from a Dutch high school (mean age = 14.31 years, $SD$ = 0.47, 25 males and 13 females).

The disparity between the number of participants in each group (i.e. 20 and 38) occurred because we did not receive consent from parents to test more 11-/12-year-olds. For all the participants in our study, we obtained written parental consent, as well as approval from their school to take part in the study. As a reward for their contribution, all participants received a small present. The current study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University. All materials regarding this study have been uploaded on the Open Science Framework and are available through the following link: http://www.osf.io/6emh2.

### *Materials*

**Pictures.** We used six photographs in total, three with a neutral and three with a negative content. The neutral images showed: (1) a classroom with two children and a teacher; (2) a cluttered desk; and (3) a kitchen with a few items on the counter. The negative pictures presented: (1) a girl in a prison cell being escorted by an officer; (2) a traffic accident involving a bicycle; and (3) a shoplifting activity in a clothes store. Each photograph contained 10 details and one critical item that was associated with the 10 details depicted on the photographs.

**DRM lists.** We used five negative DRM word lists (e.g. *manslaughter, blood, police, scary, murderer, crime, scared, perpetrator, gun, robbery*) and five neutral DRM word lists (e.g. *baker, butter, brown, dough, grain, flour, knife, wheat, old*) that were presented orally in Dutch to our participants. In total, each list contained 10 words that were semantically associated with a non-presented critical item (e.g. *murder* and *bread*). These lists have successfully been used in previous research (Howe et al., 2010).

### *Design and Procedure*

This study used a two (age group: children aged 11/12 years vs. adolescents aged 14/15 years) by two (emotional valence: neutral vs. negative) mixed factorial design, with age group as the between-subjects factor, emotional valence as the within-subjects factor, and scores on the memory conformity and DRM paradigms as dependent variables. The participants were tested in a quiet room at their schools. During the memory conformity procedure, we paired each child with a confederate of about the same age.

First, the participants looked on their own at the six photographs, which were presented on a computer screen for 20 seconds. Between each photograph, a fixation cross was shown for 4 seconds. Subsequently, the participants were instructed to remember as many details of each image as possible. They were informed that, after a distraction task (i.e. playing Tetris for 5 minutes), they would meet up with a confederate who had seen the same set of images to talk about the details of the pictures. Following this, the participants met their confederate and discussed what both could remember about the pictures, one by one. The confederates were trained by the experimenters to include false details in order to misinform the participants. More specifically, the confederates received a set of specific instructions on what they had to say during their meeting with the participants. After discussing with their confederate, the participants engaged in the distractor task again for 5 minutes, after which they were interviewed individually about what they recalled of the images. In this stage, we informed the participants that we were only interested in their own memories, not those of their confederate's.

Next, the DRM paradigm was conducted. The participants were presented with five negative and five neutral word lists, which were given to them in a counterbalanced order. After each word list, the participants were asked to recall as many words as possible. After the presentation of all four DRM word lists, a recognition task was provided that contained the 40 studied words as well as 10 critical lures, 10 related but not presented words (in addition the critical lures), and 18 unrelated words. Finally, all the participants were debriefed. For the recall task of the DRM paradigm and the memory conformity paradigm, three types of scores were obtained: false memories, true memories and intrusions. 'Intrusions' refer to words the participants reported remembering that were unrelated to the list words and the critical lures. All statistical analyses were performed using SPSS (Version 25.0).

## Results

### *Memory Conformity Paradigm*

**True memories.** A two (age group: children versus adolescents) by two (emotional valence: negative versus neutral) analysis of variance (ANOVA) was conducted on the number of true memories as measured during the final recall test. The results indicated that there was no statistically significant interaction effect of age group and emotional valence ($F(1, 56) = 0.08$, $p = 0.78$, $r = 0.04$), or statistically significant main effects

for age group ($F$(1, 56) = 0.05, $p$ = 0.82, $r$ = 0.03). Neither was there a statistically significant main effect of emotional valence on the stimuli that should be recalled ($F$(1, 56) = 1.01, $p$ = 0.32, $r$ = 0.13).

**False memories.** For the acceptance of critical items, as proposed by the confederate, during the joint discussion of the pictures, the two by two ANOVA revealed that the interaction between age group and emotional valence on the recall of false memory items was again not statistically significant ($F$(1, 56) = 0.37, $p$ = 0.55, $r$ = 0.08). There was a statistically significant main effect of age group on recalling false memories – that is, items that were falsely suggested by the confederate ($F$(1, 56) = 10.28, $p$ = 0.002, $r$ = 0.39). As shown in Table 1, 11-/12-year-old children had higher false memory rates ($M$ = 0.80, $SD$ = 0.71, 95% CI [0.56, 1.04]) than 15-/16-year-old adolescents ($M$ = 0.33, $SD$ = 0.41, 95% CI [0.16, 0.50]). The ANOVA revealed no statistically significant main effect of emotional valence ($F$(1, 56) = 2.42, $p$ = 0.13, $r$ = 0.20).

**Intrusions.** No statistically significant interaction was found between emotional valence and age on the formation of intrusions ($F$(1, 56) = 0.12, $p$ = 0.73, $r$ = 0.05). However, we found a statistically significant difference between both age groups regarding the production of intrusions ($F$(1, 56) = 8.81, $p$ = 0.004, $r$ = 0.37). The results revealed that, in general, children ($M$ = 0.88, $SD$ = 1.54, 95% CI [0.46 1.29]) developed more intrusions than adolescents ($M$ = 0.12, $SD$ = 0.27, 95% CI [–0.18, 0.42]). There was no statistically significant main effect of emotional valence on intrusions ($F$(1, 56) = 0.24, $p$ = 0.623, $r$ = 0.66).

***DRM Paradigm – Recall Test***

**True memories.** No statistically significant interaction was encountered between emotional valence and age in terms of recalling true items of the DRM lists ($F$(1, 56) = 2.83, $p$ = 0.10, $r$ = 0.22). We found a statistically significant main effect of age group ($F$(1, 56) = 9.15, $p$ = 0.004, $r$ = 37), with children ($M$ = 15.78, $SD$ = 1.72, 95% CI [15.05, 16.50]) having lower true memory scores than adolescents ($M$ = 17.13, $SD$ = 1.57, 95% CI [16.60, 17.66]). We did not find a statistically significant main effect of emotional valence with regard to the recollection of the true items in the DRM lists ($F$(1, 56) = 1.15, $p$ = 0.29, $r$ = 0.14).

**False memories.** There was no statistically significant interaction between emotional valence and age regarding the formation of false memories in the DRM recall test ($F$(1, 56) = 0.25, $p$ = 0.619, $r$ = 0.07). Furthermore, we did not find statistically significant main effects of age groups ($F$(1, 56) = 3.50, $p$ = 0.07, $r$ = 0.24) and emotion valence ($F$(1, 56) = 3.65, $p$ = 0.06, $r$ = 0.25) on the recognition of critical items in the DRM lists.

**Intrusions.** No statistically significant interaction was found between emotional valence

and age regarding the acceptance of intrusions derived from the DRM lists ($F(1, 56) = 0.78$, $p = 0.38$, $r = 0.12$). There was also no statistically significant difference between the two age groups in terms of intrusions ($F(1, 56) = 0.27$, $p = 0.61$, $r = 0.07$). There was, however, a statistically significant main effect of emotional valence on reporting items that were not related or present in the DRM lists ($F(1, 56) = 28.03$, $p < .001$, $r = 0.55$), with negative items ($M = 0.69$ $SD = 0.80$, 95% CI [0.49, 0.94]) being mentioned more than neutral ones ($M = 0.12$, $SD = 0.38$, 95% CI [0.01, 0.22]).

***DRM Paradigm – Recognition Test***

For the recognition DRM test, all scores were corrected for potential response bias. We transformed our scores using the two-high threshold correction, H – FA(U) for targets and FA(SR) – FA(U) for critical and unrelated lures (Howe & Wilkinson, 2011; Snodgrass & Corwin, 1988). H stands for the hit rates for targets, FA(SR) represents the false alarm rates for the critical lures, and FA(U) stands for the rates of unrelated lures. The correct results are presented below.

**True Memories.** There was no statistically significant interaction between emotional valence and age in terms of the recognition of correct items in the DRM lists ($F(1, 56) = 2.83$, $p = 0.10$, $r = 0.22$). We found a statistically significant age difference between both age groups regarding the recognition of true items ($F(1, 56) = 7.74$, $p = 0.007$, $r = 0.35$), in which children had lower true memory rates ($M = 0.74$, $SD = 0.11$, 95% CI [0.69, 0.78]) than adolescents ($M = 0.82$, $SD = 0.09$, 95% CI [0.78, 0.85]). In addition, no significant main effect of emotion valence was found on recognising the critical items in the DRM lists ($F(1, 56) = 1.15$ $p = 0.288$, $r = 0.14$).

**False Memories.** We found no statistically significant interaction between emotional valence and age in terms of forming false memories ($F(1, 56) = 0.25$, $p = 0.06$, $r = 0.06$). However, we did obtain a statistically significant difference between both age groups regarding the recognition of critical items ($F(1, 56) = 4.18$, $p = 0.046$, $r = 0.26$), in which children scored less ($M = 0.56$, $SD = 0.16$, 95% CI [0.47, 0.64]) than adolescents ($M = 0.66$, $SD = 0.21$, 95% CI [0.60, 0.72]), thereby demonstrating the developmental reversal effect. Moreover, no significant main effect of emotion valence was found for the recognition of critical items in the DRM lists ($F(1, 56) = 3.65$ $p = 0.06$, $r = 0.25$).

**Table 1.** Mean Scores (SDs) of DRM and Conformity Paradigm Variables Split by Age Group and Emotional Valence

| | Children | | Adolescents | |
|---|---|---|---|---|
| | Neutral | Negative | Neutral | Negative |
| Paradigms | M (SD) | M (SD) | M (SD) | M (SD) |
| Conformity | | | | |
| False memories | 0.95 (1.15) | 0.65 (0.81) | 0.40 (0.59) | 0.26 (0.50) |
| True memories | 19.50 (3.83) | 19.85 (3.60) | 19.55 (5.43) | 20.26 (4.43) |
| Intrusions | 0.95 (2.16) | 0.80 (1.51) | 0.13 (0.34) | 0.11 (0.39) |
| DRM – Recall | | | | |
| False memories | 3.25 (1.164) | 2.80 (1.24) | 3.66 (1.24) | 3.40 (1.10) |
| True memories | 15.30 (2.60) | 16.25 (2.10) | 17.24 (1.91) | 17.03 (1.82) |
| Intrusions | 0.10 (0.31) | 0.80 (0.95) | 0.13 (0.41) | 0.63 (0.71) |
| DRM – Recognition | | | | |
| False memories | 0.60 (0.24) | 0.51 (0.24) | 0.69 (0.25) | 0.64 (0.22) |
| True memories | 0.72 (0.15) | 0.76 (0.12) | 0.82 (0.11) | 0.81 (0.11) |

## Correlational Analysis

A Pearson product-moment correlation coefficient was computed to assess the relationship between the number of false memories elicited by the recall and recognition tasks of the DRM paradigm and the memory conformity paradigm. No statistically significant correlations were found for either of the relationships between false memories produced by the memory conformity paradigm and the DRM paradigm tasks (i.e. recall and recognition), as respectively presented in Figure 1 ($r$ = -.168, $n$ = 58, $p$ = .207) and Figure 2 ($r$ = -.079, $n$ = 58, $p$ = .553).

**Figure 1.** Scatterplot Showing False Memory Rates Elicited by the DRM Paradigm Recall Task and the Memory Conformity Paradigm.

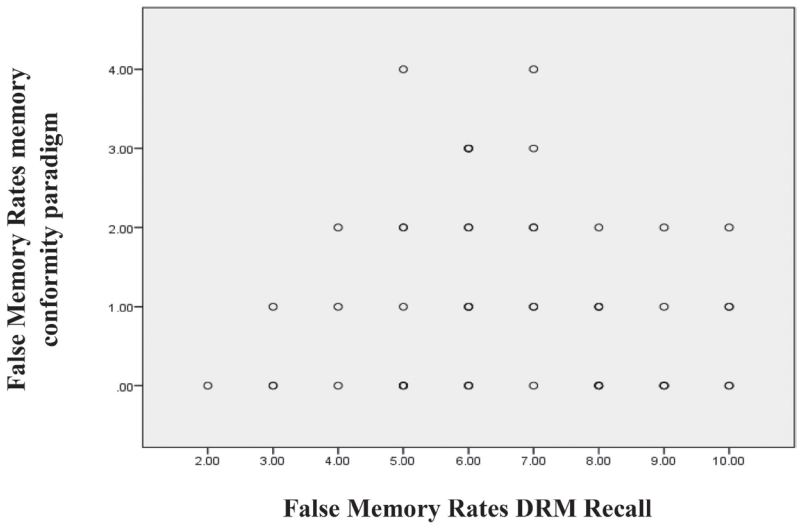
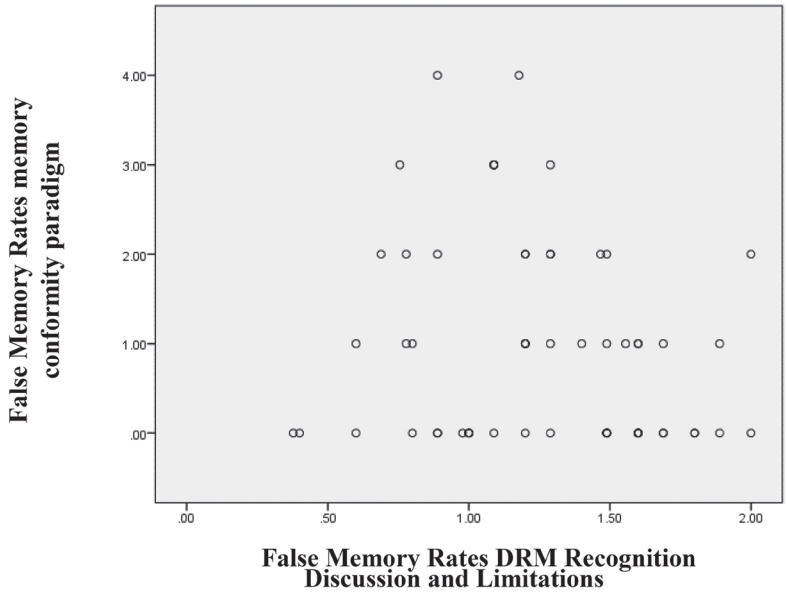


**Figure 2.** Scatterplot Showing False Memory Rates Elicited by the DRM Paradigm Recognition Task and the Memory Conformity Paradigm.








## Discussion and Limitations

The aim of this study was to investigate differences and similarities in the memory performance of older children and adolescents using two types of false memory paradigms. To be specific, we wanted to assess developmental patterns in the occurrence of false memories in a co-witness situation (the memory conformity paradigm) and in a spontaneous false memory situation (the DRM paradigm). All of our participants (i.e. older children and adolescents) were tested using a memory conformity procedure, and then consecutively conducted DRM recall and recognition tasks. We found that adolescents produced fewer false memories than children when employing the memory conformity paradigm. As expected, we found a developmental reversal when using the DRM, implying that adolescents falsely recognised more critical lures than children.

As mentioned earlier, regarding the main investigation in this chapter, developmental reversal for the formation of false memories was only statistically significant in the DRM paradigm's recognition task. Nevertheless, our results from the DRM recall task were more aligned with the DRM recognition task. Our DRM false memory results are in line with other research findings and the theoretical notions for this phenomenon, as proposed by both the fuzzy trace theory and associative activation theory. That is, the fuzzy trace theory stipulates that the capacity to retrieve gist traces from experiences increases with age, leading to a greater propensity to form false memories (Brainerd & Reyna, 2002). In addition, the associative activation theory posits that, when the associations between concepts within an individual's knowledge base are weak, fewer

false memories are formed. Furthermore, associative activation theory postulates that the association between concepts in memory become stronger with age, thereby predicting a developmental reversal effect (Howe et al., 2009).

Concerning the memory conformity paradigm, adolescents produced fewer false memories than older children. These results are in line with developmental false memory studies that used other paradigms, including suggestion (e.g. misinformation; Otgaar & Candel, 2011; Sutherland & Hayne, 2001). The memory conformity paradigm examines false memories production induced by suggestion and has been designed to mimic co-witness situations. More specifically, the memory conformity paradigm allowed us to comprehend how our participants would be influenced by a peer's input about a joint real-life experience.

Our results suggest that children are more vulnerable to their peers' suggestions than adolescents. In comparison with adolescents, children seemed to accept more easily that they could have misremembered the event they witnessed. It might also be the case that these 11-/12-year-olds are in a stage of life in which their social connections are starting to form. By agreeing with their peers, they have better chances of being accepted in potential social groups (Newman et al., 2007). Although adolescents generally share this aim to belong to their peer group, this feature might be stronger in the early stages of the transition from childhood to adolescence.

Unlike the DRM results, the results obtained using the memory conformity paradigm showed that both age groups were equally capable of remembering items that were actually presented on the photographs they had seen on the video (i.e. true items). Therefore, under appropriate circumstances (i.e. free recall), older children and adolescents are equally able to produce reliable statements from a co-witness setting. Interestingly, in the recall and recognition tasks of the DRM paradigm, the 11-/12-year-olds remembered fewer actual presented items than the adolescents. This finding indicates that, despite being less prone to memory errors under DRM recognition conditions, older children do not perform better than adolescents when it comes to remembering or recognising more details of a given crime.

Our correlational results showed no statistically significant relationship between DRM false memories (both recall and recognition) and false memories elicited by the memory conformity paradigm. This finding echoes previous work showing that false memories, as measured by different paradigms, evince a weak relationship (i.e. Bernstein et al., 2018; Ost et al., 2013; Otgaar & Candel, 2009; Patihis et al., 2013). Our results imply that spontaneous false memories and false memories induced by suggestion are formed in different ways. That is, while mechanisms such as gist extraction and associative activation are most likely to underlie the formation of spontaneous false memories, external processes such as social influences might play a more important role in the formation of suggestion-induced false memories (Brainerd et al., 2008).

This presupposition suggests that relationships between false memories will only

exist when these false memories are elicited by similar paradigms that either rely only on endogenous processes (i.e. spontaneous false memories) or exogenous processes (i.e. suggestion-induced false memories). Indeed, when Otgaar and colleagues (2013) provided children (6- to 8-year-olds and 10- to 12-year-olds) and adults with two paradigms eliciting spontaneous false memories (i.e. the standard DRM paradigm and a video variant of the DRM paradigm), statistically significant correlations did emerge between these two types of spontaneous false memories.

Our findings confirm the critical notion that the term 'false memory' should not be used universally, since different paradigms elicit different types of false memories (i.e. Patihis et al., 2018). Thus, studies in the literature advocate for the importance of clarifying the term 'false memories', which seems to be particularly relevant for legal cases. For example, Patihis and colleagues (2018) described a legal situation in which it was assumed that, because a victim of sexual abuse had scored low on a suggestibility test, she would not be prone to false memories in general. Contexts such as this case show that it is of utmost importance to clarify these differences among the mechanisms for different types of false memories in the literature, since research on false memories is used as a valued source of information by legal practitioners. The fact that we were unable to document correlations between the false memories elicited by two different paradigms supports the idea that the term 'false memory' is not unitary, and that different mechanisms might underlie these false memories.

In our study, negative emotional valence did not seem to play an important role in the formation of false memories in either of the paradigms. The exceptions to this finding were the intrusion rates for the DRM paradigm recall test. In this case, our results showed that negative words were more prominent during free recall than neutral words. These results are not in line with the reported findings in earlier work (see Bookbinder & Brainerd, 2016; Talmi et al., 2007).

According to Otgaar and colleagues (2009), for instance, 7-year-olds heard two true stories and one false story about their school experiences. Half of the participants received a neutral false story (that they would be moving to a different classroom), while the other half received a negative false story (that they were being accused of cheating by their teacher). The negative false story provoked much higher levels of false memories than the neutral one. In the current study, our negative photographs in the memory conformity paradigm may not have been sufficiently emotional-laden, which may have decreased the participants' emotional perception of the DRM paradigm word lists (the subsequent task). Considering that the pictures presented more details than the words on the DRM lists, if the participants did not find the pictures negative, they would certainly not find the words on the list negative either, if they corresponded to the same theme as the photographs.

One limitation in our study is related to the disparity in the number of participants in both examined groups (i.e. 11-/12-year-olds and 13/14-year-olds). That is, the

number of 13-/14-year-olds that we tested (38) was almost twice the number of tested 11-/12-year-olds (20). Hence, it is possible that our study suffered from a power problem, and future research should attempt to replicate our results in larger samples.

## Conclusion

The increasing number of children being brought into the legal arena, as both eyewitnesses and victims, has fuelled academic and legal discussions concerning children's ability to accurately recall and report crimes they have experienced (Bruck & Ceci, 1999; Flin et al., 1992). The present study contributes to this discussion by comparing older children's and adolescents' vulnerability to false memories in two different paradigms that relate to different real-life contexts. We found considerable false memory differences between both groups (11-/12-year-olds and 14-/15-year-olds) in the current study and observed no notable relation between different types of false memory. Our results imply that researchers should continue to design studies with elements that simulate real life in order to obtain a better grasp of how false memories due to suggestion or spontaneous false memories might appear in real-life cases such as child sexual abuse.

# Chapter IV

## Statements of Sexual Abuse Revictimisation are Assessed as Less Credible than Statements of Single-Occurrence Sexual Abuse

Calado, B., Luke, T. J., Landström, S., Zahn, P., & Otgaar, H.

*This chapter corresponds to the following paper:*
Statements of Sexual Abuse Revictimization Are Assessed as Less Credible Compared to Statements of a Single-Occurrence Sexual Abuse.
Submitted to: Journal of Interpersonal Violence.

**Abstract**

Physical evidence and corroborating witness statements are often dubious or absent in cases of sexual abuse revictimisation. Therefore, victims' testimonies are a crucial piece of evidence for legal decision-making. Jurors and practitioners of the legal system rely heavily on the credibility they attribute to the statements of alleged victims of sexual abuse revictimisation in their decision-making process. In the current experiment, practitioners of the legal system (i.e. police officers, social workers, lawyers, judges, and prosecutors), law students, and the general public were asked to evaluate the credibility of statements of sexual abuse victims describing revictimisation by the same perpetrator and of victims reporting single-occurrence abuse. Furthermore, we manipulated the victim characteristics in terms of developmental group and time of disclosure (i.e. a child reporting current sexual abuse, an adult reporting current sexual abuse, and a child reporting historical abuse), but found that this did not interfere with the participants' perceived credibility of the victims' statements. The participants did, however, attribute less credibility to the victims' statements that contained descriptions of multiple accounts of sexual abuse by the same perpetrator under similar conditions as victims who reported one-time abuse. Our results are aligned with previous research findings.

***Keywords:*** sexual abuse, credibility, revictimisation, repeated events

## Opening

In the 1980s and 1990s, thousands of alleged victims in the United States and in some European countries reported memories of being repeatedly sexually abused as a child, after attending suggestive therapy sessions (for reviews, see Nathan & Snedeker, 2001; Young, 2008; see also Goodman et al., 1997). Moreover, hundreds of children gave ritualistic and bizarre accounts of repeated sexual abuse allegedly perpetrated by several teachers in their day care centres, after being exposed to suggestive interviews and peer influences (Garven et al., 1998). Since then, psychologists have increasingly studied memory's malleability and its implication in the courtroom as a means of explaining these improbable and implausible accounts (Loftus, 2005; Loftus & Pickrell, 1995). This research has clearly shown that people can form false autobiographical memories (Scoboria et al., 2017).

## Child Sexual Abuse Cases and Revictimisation

False memories can refer to an event that happened only once, or to repeated occurrences of the same event (see Chapter V). This is particularly relevant in sexual abuse cases, since about two-thirds of sexual abuse victims are revictimised by the same perpetrator (Classen et al., 2005; Messman-Moore & Long, 2003). The notion of false memory formation in the legal system has reinforced the uncertainty expressed by triers of fact regarding the veracity of sexual abuse statements, particularly when DNA traces and other forms of physical evidence are absent (Albright, 2017).

In addition to the fallibility of memory, the frequency with which an alleged victim experienced abuse might influence how the legal system attributes credibility to sexual abuse allegations (Weinsheimer et al., 2017). For instance, children who describe experiencing a repeated event may be perceived as less credible than children who narrated experiencing a single event (Connolly et al., 2008[b]). In Connolly and colleagues (2008), evaluators watched videos of children providing memory reports of play sessions. Children narrated either a unique play session or multiple play sessions. Children who experienced repeated play sessions showed some variations in the details they gave of each play session occurrence, despite the script of the play session being the same. In Connolly and Lavoie (2015), in addition to reporting actual events, children also provided a report about a fabricated event that had (supposedly) happened either once or multiple times. The results of both studies were generally consistent, with evaluators assessing children who described repeated events as less credible than children who provided a memory report about a unique event.

## Credibility of Sexual Abuse Victims Based on Age

Besides the frequency of the reported events, the victim's age and the time of disclosure might be relevant features affecting the perceived credibility of a sexual abuse statement. The population of sexual abuse victims comprises both children and adults, with children representing a large part of this population. At least 1 in 10 children will be sexually abused before the age of 18 (Child Sexual Abuse Statistics, 2018); furthermore, in some countries, children make up the higher percentage of victims of sexual abuse. In Brazil, for example, children and adolescents comprise 70% of the population of sexual abuse victims (70% das Vitimas, 2017).

Problematically, children may be perceived as incompetent when it comes to producing a reliable legal statement, particularly when compared with adults (Brainerd, 2013; Bruck & Ceci, 1999). In fact, when being questioned in a suggestive manner during an investigative interview, adults are less prone to fall prey to suggestion, as compared with children (Ceci & Bruck, 1993). However, children can perform more competently than adults in specific situations. For instance, children are less likely than adults to make errors when freely recalling past episodes (Otgaar et al., 2018). Moreover, among sexual abuse cases, a non-trivial number of sexually abused children never report the crime (Fisher et al., 2003). Some child sexual abuse victims only make a police report during their adulthood, which qualifies as historical sexual abuse (Shead, 2014). It is unclear whether historical sexual abuse reports are given the same credibility as statements that narrate a current memory report of sexual abuse.

## The Current Study

Because physical evidence is commonly absent in sexual abuse cases, particularly in historical sexual abuse, the credibility assessment made by legal practitioners may be crucial for the case's verdict (Albright, 2017). Legal professionals are not, however, the only parties who decide on the verdict of a legal case. Certain trial configurations demand a sworn body of laypeople to render a judgment: the jury. Necessarily composed of the general public, the jury is an important piece in the legal system of countries with a common law system (Hans, 2008). Besides the possibility of being summoned to serve as a juror, civilians can affect legal cases via their influence over the media. When the population of a society pressures the media to consistently follow up on a given case, the justice process may be speeded up as a result (Casillas et al., 2010).

In the current experiment, we investigated whether the sexual abuse statements of victims who reported revictimisation would be considered less credible than the statements of victims who allegedly experienced one-time sexual abuse. Furthermore, we wanted to understand whether the credibility assessment of legal practitioners (i.e. police officers, lawyers, judges etc.), law students, and the general public would vary if

the victim was (1) a child who reports current abuse; (2) an adult who reports current abuse; or (3) an adult who reports historical sexual abuse. Hence, each participant was asked to assess the credibility of a single-occurrence and a multiple-occurrence sexual abuse statement within one of three victim profiles: 'child current abuse', 'adult current abuse', and 'adult historical abuse'.

We made the following hypotheses. (1) The participants would rate reports of single sexual abuse events as more credible than reports of repeated sexual abuse occurrences. (2) Adults reporting repeated historical sexual abuse would be rated as less credible than adults reporting one-time historical sexual abuse. In addition, we hypothesised that the difference between single versus multiple sexual abuse occurrences would be smaller in the child current abuse cases than in the adult historical abuse cases. (3) Statements from child current abuse victims would be rated as less credible compared with adult current abuse victims. (4) Adults who reported historical sexual abuse would be rated as less credible than children reporting current sexual abuse. (5) When comparing adult historical abuse and adult current abuse cases, adult victims reporting current sexual abuse would be deemed more credible than those describing historical sexual abuse. (6) Furthermore, we wanted to assess in an exploratory manner whether there would be a difference in credibility evaluation among the three types of sample in the current study (i.e. practitioners of the legal system, law students, and the general public).

## Method



The Ethics Review Committee of Psychology and Neuroscience (ERCPN) of Maastricht University approved the current experiment. We preregistered the method and made the materials for the present study available on the Open Science Framework (OSF). Before data collection was completed and prior to data analysis, we updated the registration with a more specific analysis plan: https://osf.io/kauzc/?view_only=1f5b55d5791d460abdaa91fee015fc05.

### *Participants*

An *a priori* power analysis using G*Power (Erdfelder et al., 1996) indicated that a total of 129 participants were needed for each of the three samples (practitioners, law students, and the general public); that is, with 95% power, there was a medium to large effect size ($f$ = .30) and a default correlation among repeated measures ($r$ = .50). Each experimental group, named after the cases to be assessed (i.e. child, adult, and adult historical) had a minimum of 43 participants from each sample (i.e. law students, practitioners of the legal system, and members of the general public). We based the effect size used in this power analysis on previous studies in which credibility was assessed via vignette studies (i.e.Connolly et al., 2008[b]; van Veldhuizen, et al., 2017; Tenney et al., 2008).

We excluded participants based on two criteria: first, we excluded subjects that failed

to answer all or some of the questions in our study. Second, we excluded participants who omitted their profession or age. For the sample of practitioners of the legal system, we also excluded participants who reported working in a profession outside of the legal field, or who reported being retired.

***General Public***

Our general public sample was composed of individuals registered on the Mechanical Turk Amazon platform. In total, we recruited 243 participants; after applying the exclusion criteria, we ended up with 151 participants (mean age = 34.64 years, *SD* = 10.66, 90 males and 61 females). The participants received $1 USD for their participation.

***Law Students***

We recruited a total of 310 law students from a Belgian university (*n* = 150), a German university (*n* = 154), and a Spanish university (*n* = 6). After applying the exclusion criteria, we ended up with 187 participants (mean age = 20.58 years, *SD* = 0.24, 39 males and 148 females). Additional demographic information about the law student sample can be found in Appendix A.

***Practitioners of the legal system***

Practitioners of the legal system were recruited through their work institutions and personal contacts. This resulted in the acquisition of participants from around the world. The study was created and distributed online by sending and sharing the corresponding link through email or social media (e.g. Facebook). Participation was fully voluntary and did not include any compensation. In total, we recruited 337 participants that professionally work within the legal arena. After applying the exclusion criteria, we ended up with 221 participants (mean age = 42.91 years, *SD* = 13.67, 109 males and 112 females) residing in Brazil (*n* = 15), Canada (*n* = 1), Denmark (*n* = 1), Finland (*n* = 77), France (*n* = 1), Germany (*n* = 51), Iceland (*n* = 1), Ireland (*n* = 3), the Netherlands (*n* = 1), Spain (*n* = 8), Sweden (*n* = 50), Switzerland (*n* = 1), the United Kingdom (*n* = 9), and the United States (*n* = 2). The participants reported different professions, such as police officers, judges, and lawyers who interview and assess the credibility of statements from victims and eyewitnesses as part of their profession. Additional demographic information about the legal system practitioner sample can be found in Appendix A.

***Design***

We used a three (*victim*: child, adult, and adult historical) by two (*frequency*: single vs. multiple) mixed design, with *victim* being between subjects and *frequency* being within subjects. The frequency order (i.e. whether the participant saw the single abuse vignette or the multiple abuse vignette first) was allocated to each participant in a counterbalanced fashion.

### *Materials*

We used the online survey distribution program Qualtrics to present the participants with the fictional vignettes containing sexual abuse statements, questions related to the vignettes and a questionnaire containing demographic questions. All materials were translated into and available in eight different languages (English, German, Finnish, Swedish, Dutch, French, Spanish, and Portuguese).

### *Vignettes*

The vignettes (see Appendix A) presented to the participants were three fictional stories about sexual abuse. The settings of the alleged abuse included a bus, a school bathroom, and the basement of a friend's house. Due to the high prevalence of female victims in cases of sexual abuse, we chose a female victim and a male offender for our vignettes (RAINN, 2019; The National Centre for Victims of Crime, 2011). Each fictional story was adapted into a single and a multiple frequency version. Each participant saw one story in its single format and another story in its multiple format. We randomised the stories in such a way that the participants would never see the same story in its single and multiple versions. In the single vignettes, the word variation ranged from 79 to 101 words; in the multiple vignettes, the word variation ranged from 142 to 172 words. We asked the participants to read and evaluate both stories independently in terms of credibility on a 6-point Likert scale: 1 – 'extremely unlikely', 2 – 'unlikely', 3 – 'somewhat unlikely', 4 – 'somewhat credible', 5 – 'credible', and 6 – 'extremely credible'. The participant was then asked (in an open-ended way) to state which of the two stories they deemed to be more credible and why.

### *Demographics Questionnaire*

The demographics questionnaire (see Appendix B) consisted of 10 demographic questions that included questions about participants' sex (male or female), gender (prefer not to state, male, female or other), age, nationality, ethnicity, country, profession (for students, we asked what year of their degree they were in), and current occupation, years spent working in that profession, educational background, specialisation, and side job.

### *Procedure*

The study was distributed by sharing the corresponding Qualtrics link through social media (e.g. Facebook) or emailing it to the participants. After reading the information about the study and signing a declaration of consent online, the experiment began with the presentation of the sexual abuse statements. First, information about the origins of these statements was displayed – that is, the participants were informed that the statements came from one of the following sources: a 5- to 8-year-old female child (child current abuse), a 21-year-old female adult whose sexual assault had happened within the last 3 years (adult current abuse) or a 21-year-old female adult whose sexual assault

had happened during childhood (adult historical abuse). Subsequently, the participants viewed two vignettes describing sexual abuse: one described a single-occurrence abuse and the other described a multiple-occurrence abuse by the same perpetrator. After reading each vignette, the participants were prompted to assess the credibility of the vignettes on the 6-point Likert scale. After the participants rated both the vignettes, in order to assess whether the participants were conscious of how they rated both vignettes, the participants were asked to indicate which of the two vignettes they had considered more credible and to explain why. After the experiment, the participants filled a demographic questionnaire with 11 items. In the last stage of the experiment, the participants were debriefed, and the study's outlined rationale, hypotheses, and design were presented to the participants.

## Results

The current study aimed to investigate: how (1) the frequency of sexual abuse occurrence (by the same perpetrator), and (2) the victim's age (child vs. adult) combined with (3) the time of disclosure (current abuse vs. historical abuse) affect the credibility assessment of sexual abuse statements. We formed our complete sample from three different populations: the general public, law students, and legal professionals.

After the participants rated the two vignettes (one single-occurrence and one multiple-occurrence), we asked them to assess which of the two statements they found more credible. In addition, we requested the participants to freely write the reasoning behind their evaluation. We then conducted a content analysis of the practitioners of the legal system's responses to open-ended questions (Bardin, 2011). We did not code the answers to the open-ended question from the general public and law students, as over 50% of the participants in both of these samples provided unclear or ambiguous answers, such as 'the other one was more believable' or 'more details' without specifying the direction. Approximately 30% of the participants in both samples also chose not to respond why they had made their choice.

### Credibility Experiment

Table 1 provides a descriptive analysis of the participants' credibility ratings for a child victim reporting current sexual abuse, an adult victim reporting current sexual abuse and an adult victim reporting historical sexual abuse. For each of the three conditions (*victim*, *age*, and *time of disclosure*), there is a statement describing sexual abuse as a single occurrence and another statement describing a multiple-occurrence abuse.

**Table 1.** Descriptive Statistics for Sexual Abuse Victims' Credibility

| Sexual abuse frequency | Victim | Sample | *N* | Mean | *SD* | Median |
|---|---|---|---|---|---|---|
| Single occurrence | Child current sexual abuse occurrence | General public | 53 | 4.36 | 1.09 | 5 |
| | | Law students | 63 | 4.18 | 1.02 | 4 |
| | | Practitioners of the legal system | 73 | 3.86 | 1.02 | 4 |
| | Adult current sexual abuse occurrence | General public | 48 | 4.21 | 1.35 | 4 |
| | | Law students | 63 | 4.23 | 1.09 | 4 |
| | | Practitioners of the legal system | 74 | 3.88 | 1.16 | 4 |
| | Adult historical sexual abuse occurrence | General public | 50 | 4.62 | 1.21 | 5 |
| | | Law students | 61 | 3.84 | 0.96 | 4 |
| | | Practitioners of the legal system | 74 | 3.54 | 1.25 | 4 |
| Multiple occurrences | Child current sexual abuse occurrence | General public | 53 | 4.11 | 1.49 | 5 |
| | | Law students | 63 | 3.88 | 1.11 | 4 |
| | | Practitioners of the legal system | 73 | 3.48 | 1.27 | 4 |
| | Adult current sexual abuse occurrence | General public | 48 | 3.94 | 1.56 | 4 |
| | | Law students | 63 | 3.68 | 1.28 | 4 |
| | | Practitioners of the legal system | 74 | 3.24 | 1.34 | 3 |
| | Adult historical sexual abuse occurrence | General public | 53 | 4.20 | 1.29 | 4 |
| | | Law students | 62 | 3.57 | 1.16 | 4 |
| | | Practitioners of the legal system | 73 | 3.72 | 1.26 | 4 |



**Analytic Approach.** Due to the nested nature of our data, we used linear mixed-effects modelling to assess the differences in the credibility attributed to each condition of our experiment (child current abuse, adult current abuse, and adult historical abuse). We used the lme4 package (Bates et al., 2015) for R. To obtain Satterthwaite approximated degrees of freedom, we complemented the lme4 package with the lmerTest package (Kuznetsova et al., 2017). Our models used restricted maximum likelihood (REML) estimation. Each participant provided up to two measurements, one for each statement they assessed (single vs. multiple). In total, we analysed $N = 1113$ observations.

**Model Selection.** We selected the appropriate model using a hierarchical approach. The first step consisted of fitting a model that only included the main effects of each of the fixed factors. In the second step, we added the two-way interactions. We compared both models using a likelihood ratio test. The comparison required the models to be refitted using maximum likelihood instead of REML; the results we report correspond to the REML fittings. In assessing Hypotheses 1–5 (all considering the general sample

*N* = 558), the second model that included the interactions presented a significant improvement to the data in comparison with the first model ($\chi^2$ (2) = 6.83, $p$ = .03). However, none of the interaction coefficients were different from 0 in the second model; hence, we retained the first model containing the main effects only.

To assess Hypothesis 6, which proposed a comparison of the samples (general public, law students, and practitioners of the legal system), we added the sample variable to the third model. The fourth model, which included two-way interactions, did not fit the data significantly better than the third model that included the main effects only ($\chi^2$ (4) = 5.90, $p$ = .42). Hence, we retained the third model. As the coefficients in the first and third models for the main effects of the general sample ($N$ = 558) were almost identical, we simply retained the third model to explain all of our six hypotheses. We report the results for the third model here; however, the other models are available in our supplementary material (see https://osf.io/kauzc/?view_only=1f5b55d5791d460abdaa91fee015fc05).

To assess Hypotheses 1–6, the model included fixed effects for *victim/time of disclosure* (child current abuse, adult current abuse or adult historical abuse), *frequency* (single occurrence or multiple occurrence), and random intercepts for participants. Specifically in order to test the sixth research question, in addition to fixed effects for *victim/time of disclosure* and *frequency* (single occurrence and multiple occurrence), we included fixed effects for the sample (general public, law students, and practitioners of the legal system). We added fixed effects for the sample in this case because Hypothesis 6 questions whether there is a difference among the three samples in terms of the credibility attributed to sex abuse victims.

**Modelling Results**

**Hypothesis 1.** As displayed in Table 2, the results supported our prediction that our sample would attribute higher rates of credibility to sexual abuse reports depicting single-occurrence abuse compared with reports describing multiple-occurrence abuse.

**Hypothesis 2.** The difference in credibility rating between single-occurrence and multiple-occurrence sexual abuse was not higher in the adult historical abuse cases compared with the child current abuse cases. Hence, Hypothesis 2 was not supported.

**Hypothesis 3.** We predicted that less credibility would be attributed to the children's statements by the participants viewing those vignettes than the credibility attributed to the adult current abuse cases by the participants viewing those vignettes. That is, regardless of abuse frequency, we believed that children would be found to be less credible than adults when reporting (non-historical) sexual abuse occurrences. Our data do not support this hypothesis. There was no significant difference between the child current abuse and the adult current abuse cases in this study across the three samples.

**Hypotheses 4 and 5**. We predicted that the participants would find adult victims reporting historical abuse in either frequency (single, repeated) to be less credible than a child victim reporting current abuse. We did not find support for this hypothesis. Furthermore, we predicted that adults reporting current sexual abuse would be found to be more credible by the participants than adults reporting historical abuse. Again, this hypothesis was not supported by our data. There was no statistically significant difference between the three experimental groups in our study in terms of statement credibility assessment.

**Hypothesis 6.** Finally, our last (exploratory) hypothesis questioned whether there would be a difference between the three samples in terms of the credibility ratings the participants attributed to sexual abuse statements. Our data show that there was a statistically significant difference between the three samples in terms of their credibility ratings. Law students and practitioners rated the sexual abuse statements as less credible than the general public sample did. Practitioners attributed even lower ratings than law students.

**Table 2.** Linear Mixed-Effects Model Results for the Credibility of Victims.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***t-value*** | ***df*** | ***p value*** |
|---|---|---|---|---|---|
| Intercept | 4.45 [4.23, 4.67] | 0.11 | 38.22 | 44.18 | <.001 |
| Multiple occurrences | -0.31 [-0.41, -0.21] | 0.05 | -5.54 | 554.03 | <.001 |
| Adult current abuse | -0.10 [-0.30, 0.10] | 0.10 | -1.02 | 555.73 | .30 |
| Adult historical abuse | -0.08 [0.28, 0.12] | 0.10 | -0.76 | 555.73 | .44 |
| Law students | -0.33 [-0.55, -0.11] | 0.11 | -3.01 | 556.23 | .00266 |
| Practitioners of the legal system | -0.61 [-0.81, -0.41] | 0.10 | -5.76 | 554.71 | <.001 |
| Random effects | SD | | | | |
| Participants | 0.79 | | | | |
| Victim condition | 0.08 | | | | |

Note: Coefficients are displayed with 95% CIs.

**Content Analysis.** We coded the 221 answers to the open-ended question we collected from the practitioners of the legal system. In Table A1, Table A2 and Table A3, we present the final categories from the content analysis (Bardin, 1977) conducted by the first author of the current study and an independent judge. Each table corresponds to the results for one experimental group (child current sexual abuse cases, adults current sexual abuse cases, and adult historical sexual abuse cases) from the legal system practitioner sample only. We present the results in forms of counts on each category, with the participants' ID numbers under the corresponding categories matching the participants' answers. The final categories are included in Table 3. Importantly, these categories are not mutually exclusive, so one response could be assigned to more than one category.

**Table 3.** List of Categories Drawn from Legal Practitioners' Justification for the Credibility They Assigned to Sexual Abuse Statements.

| **Categories** |
|---|
| 1- More details make the statement more credible. |
| 2- The abuse context seemed more credible. |
| 3- Repeated accounts of abuse are less credible. |
| 4- The victim's behaviour did not seem credible. |
| 5- The proximity with the aggressor makes the statement more credible. |
| 6- The language used in the statement does not match a child (specific to the child current sexual abuse group). |
| 7- Unclear answer, or chose not to respond, or stated being unable to choose between either statement. |
| 8- Both statements equally credible. |
| 9- Description of victim's feelings makes the statement more credible. |
| 10- Simpler statements are more credible. |
| 11- Preconceptions on how a victim should report sexual abuse occurrences. |

For the current investigation, Category 3, 'Repeated accounts of abuse are less credible' is the most relevant category that emerged from the participants' answers. One of our main interests was to assess differences in how credibility was attributed to single-occurrence versus multiple-occurrence accounts of sexual abuse. In total, 28 (12.6%) answers from the practitioners of the legal system sample fit into Category 3. In Category 3, the participants chose the single-occurrence statement as more credible than the multiple-occurrence statement precisely because a statement reporting repeated episodes of sexual abuse did not sound reliable to them.

Category 1's high frequency of 61 answers indicates that 27.6% of the practitioners of the legal system who participated in our study attributed the number of details in a sexual abuse disclose statement as an important factor in crediting reliability to the victim's claims. However, the participants' answers in this category did not indicate the particular details that shaped their decision. The participants did, however, indicate that more details increased the statement's credibility. Category 7 also had a high frequency of 61 answers (27.6%). This category, which corresponds to 'Unclear answer, chose not to respond or stated not to be able to choose between either statements', might indicate that the practitioners' considered themselves to lack the necessary information to assess such cases.

Here, it is relevant to note that Category 8, in which the practitioners considered both statements to be equally credible, had a low frequency of 14 answers (6.3%), whereas Category 4, 'The victim's behaviour did not seem credible', had a higher frequency of 27 answers (12.2%). The practitioners' answers in Category 4 often involved literal judgments of how the victims behaved in the statements. As a complement to Category 4, Category 11, 'preconceptions on how a victim should report sexual abuse occurrences', had a frequency of 11 practitioners (4.9%).

## Discussion

The aim of this study was to assess how practitioners of the legal system, law students, and the general public perceived the credibility of a statement containing repeated accounts of sexual abuse versus a statement depicting a single occurrence of sexual abuse. Furthermore, we investigated whether this credibility assessment would vary if the victim was a child disclosing current sexual abuse, an adult disclosing current sexual abuse or an adult disclosing historical sexual abuse. The participants in this study attributed significantly less credibility to statements describing repeated claims of sexual abuse, compared with statements narrating a single-occurrence sexual abuse. Hence, our first hypothesis was supported. These results are in line with previous studies investigating how university students assessed the credibility of memory reports of single versus repeated events.

Weinsheimer and colleagues (2017) found that memory statements describing food tasting experiences were deemed to have low credibility when the report included multiple occurrences. Connolly et al. (2008[b]) found similar results; however, in that case, the credibility assessment were given only about children. With the exception of Hypothesis 1, none of the other hypotheses were supported by our data. That is, we did not find any significant difference in credibility being attributed to child versus adult victims, nor did we find significant differences in credibility being attributed to adult victims reporting current sexual abuse versus historical sexual abuse. In addition, our data suggested that practitioners of the legal system and law students attributed less credibility overall to all victims' statements, compared with the general public.

The content analysis we conducted on the practitioners of the legal system' answers suggested that it is possible that the participants did not perceive our manipulation. This was particularly the case for the question that required the participants to justify why they chose one of the sexual abuse statements to be more credible than the other statement (single versus multiple occurrences). A substantial number of responses suggested that the participants based their judgment on the number of details provided in the statements they assessed. The practitioners did not expand on what kind of details they were referring to, which made our interpretation more difficult, since we fashioned the vignettes to be similar in terms of their main elements (female victim, non-family member perpetrator) and length.

Furthermore, the practitioners of the legal system explained that the reports containing repeated sexual abuse occurrences did not sound credible because, once having been abused, the victim would definitely avoid the aggressor or avoid being in a context similar to that of the first abuse. The practitioners also found it unlikely that the victims who suffered sexual abuse repeatedly would not report the abuse to the police or to friends and family. Of the 11 categories based on the practitioner's answers, 10 categories reinforced stereotypes of how sexual abuse victims 'ought' to behave and

communicate their abuse (see Kennedy & Prock, 2018; Randall, 2010). It certainly seems alarming that 10 out of the 11 categories that emerged from the statements of the practitioners' sample in our study indicated a reliance on stereotypes of victim behaviour, rather than research-based accounts of how victims tend to behave.

Since two-thirds of the victims of sexual abuse experience revictimisation (Classen et al., 2005; Messman-Moore & Long, 2003), the indication that people perceive multiple-occurrence sexual abuse statements as less credible compared with single-occurrence sexual abuse statements can be considered to be a significant problem. Such indications of the prejudices that are held about what a 'true' sexual abuse statement looks like is particularly worrying in the case of practitioners of the legal system, who are directly involved in the investigation and prosecution of such cases. For instance, prosecutors are likely to make use of rape stereotypes, such as what a rape scenario 'should' look like, or even what the 'correct' timeframe is to report sexual abuse (Spohn et al., 2001; Frohmann, 1991). Spohn and colleagues (2001) looked into the reasoning behind case rejections by prosecutors in the United States and found that, when making decisions, prosecutors were most concerned about how likely a conviction in court would be. Prosecutors based their assessment on stereotypes of what constitutes a 'true' rape victim (Spohn et al., 2001). Hence, it is likely that even trained legal practitioners judge cases based on stereotypes, rather than on objective credibility criteria.

***Limitations***

When designing the vignettes in the current experiment, we wanted to minimise factors we were not interested in measuring as much as possible. In order to solely assess the issue of sexual abuse victims' statement credibility for single-occurrence versus multiple-occurrence incidents, we levelled the language for all three types of victims in our study. That is, the level of communication displayed by the child and adult victims was the same. However, some of the participants that viewed the child victim cases explained that they did not believe a child had the knowledge to use a certain level of vocabulary. Thus, future research could adapt the children's statements to use language appropriate to their developmental stage.

## Conclusions

Our experiment shows that victims who suffer sexual abuse revictimisation are more likely to have their credibility jeopardised. The general public, law students, and – importantly – practitioners of the legal system deemed sexual abuse statements containing multiple-occurrence abuse to be less credible than single-occurrence abuse. Such a potential bias towards victims of repeated sexual abuse heightens the risk of victim-blaming behaviours. Furthermore, if cases of repeated sexual abuse are considered to have low credibility, such cases are less likely to be prosecuted. Making practitioners

of the legal system aware of this potential bias could ensure more sensitive and equal treatment of sexual abuse victims in the criminal justice system. In terms of credibility assessment, the legal system should be better equipped in operationalising the relatively few sexual abuse cases that are reported in order to avoid increasing the victims' trauma and potentially decreasing their trust in the legal system.

## Appendix A

***Introduction Text to the Vignettes***

**Child Current Abuse Group**

Next, you will read 2 short stories of sexual abuse allegations that originate from female children between the ages of 5 and 8. All stories are reported to have happened at least one year prior to the given accounts, and there is no corroborating physical evidence to support or discredit the alleged victims' statements. In addition, the accused denies all allegations.

**Adult Current Abuse Group**

Next, you will read 2 short stories of sexual abuse allegations that originate from 21-year-old females. All stories are reported to have happened within the time frame of three years prior to the given accounts, and there is no corroborating physical evidence to support or discredit the alleged victims' statements. In addition, the accused denies all allegations.

**Adult Historical Abuse Group**

Next, you will read 2 short stories of sexual abuse allegations that originate from 21-year-old females. All stories are reported to have happened during the victim's childhood, between the ages of 5 and 6, and there is no corroborating physical evidence to support or discredit the alleged victims' statements. In addition, the accused denies all allegations.

**Vignette 1 – Single-Occurrence Abuse**

*It was my second week of classes; I was on the bus that takes me from school to my house. At some point, I realised that it was just me and the driver left in the bus. I didn't know where he had taken us. He stopped the bus, came to the back seats and penetrated me. He was hurting me. He was way too heavy and was moving. I was crying. He got out of me and came on my stomach; I was so confused. He said I shouldn't mention this to anyone, if I did, he would kill me.*

**Vignette 1 – Multiple-Occurrence Abuse**

*It was my second week of classes; I was on the bus that takes me from school to my house. At some point I realised that it was just me and the driver left in the bus. I didn't know where he had taken us. He stopped the bus, came to the back seats and penetrated me. He was hurting me. He was way too heavy and was moving. I was crying. He got out of me and came on my stomach; I was so confused. He said I shouldn't mention this to anyone, if I did, he would kill me. It happened again many times until I changed the bus company. The last time it happened I remember I was already so afraid, I felt it coming. It was the same as past times; I was the last one on the bus, and despite getting out quickly he followed me to the street and dragged me to a construction site. He threatened my life, took off my clothes and raped me.*

**Vignette 2 – Single-Occurrence Abuse**

*He was in my Spanish classes; he was much older than me. The class had finished, and I had gone to the bathroom. When I opened the door, he was just outside of the door, close to the sink. Confused, I asked him what he was doing there in the girls' bathroom. He then pushed me back inside the toilet cabin and started to lift my dress. He penetrated me, I was telling him to stop, but he wouldn't.*

**Vignette 2 – Multiple-Occurrence Abuse**

*He was in my Spanish classes; he was much older than me. The class had finished, and I had gone to the bathroom. When I opened the door, he was just outside of the door, close to the sink. Confused, I asked him what he was doing there in the girls' bathroom. He then pushed me back inside the toilet cabin and started to lift my dress. He penetrated me, I was telling him to stop, but he wouldn't. It happened many times before I changed my schedule. The last time it happened he followed me home. I realised someone was walking behind me; when I saw him, he dragged me to a construction site, telling me not to scream. He threatened my life and raped me.*

**Vignette 3 – Single-Occurrence Abuse**

*It was during a lovely Christmas party my friend and her family were hosting. Her father has always been a little inappropriate with her friends. I just never though he would be capable of doing this. He asked me to help him to carry some things to the basement. When we got there, he told me not to scream, or else he would kill me. He asked me to take off my clothes, took off his trousers and raped me.*

**Vignette 3 – Multiple-Occurrence Abuse**

*It was during a lovely Christmas party my friend and her family were hosting. Her father has always been a little inappropriate with her friends. I just never thought he would be capable of doing this. He asked me to help him to carry some things to the basement. When we got there, he told me not to scream, or else he would kill me. He asked me to take off my clothes, took off his trousers and raped me. It happened many times until I stopped going to my friend's house. The last time it happened was at my friend's birthday party. I was cleaning up a few things in the kitchen when he approached me from behind. He told me not to do anything and threatened my life. I heard the zipper of his trousers opening, then he raped me.*

# Additional Demographic Information

## *Law Students Sample*

**Table A1.** Law Students' Declared Country of Origin.

| Country of Origin | Frequency | (%) |
|---|---|---|
| Belgium | 68 | 36.4 |
| Finland | 1 | .5 |
| France | 1 | .5 |
| Germany | 93 | 49.7 |
| Greece | 1 | .5 |
| India | 1 | .5 |
| Italy | 2 | 1.1 |
| Mexico | 1 | .5 |
| Netherlands | 6 | 3.2 |
| Spain | 8 | 4.3 |
| Turkey | 1 | .5 |
| United Kingdom | 4 | 2.1 |
| **Total** | **187** | **100** |

**Table A2.** Law Students' Declared School Year.

| Year of Undergraduate Studies | Frequency | (%) |
|---|---|---|
| 1 | 5 | 2.7 |
| 2 | 76 | 40.6 |
| 3 | 41 | 21.9 |
| 4 | 57 | 30.5 |
| 5 | 4 | 2.1 |
| 6 | 2 | 1.1 |
| Missing | 2 | 1.1 |
| **Total** | **187** | **100** |

## *Practitioners of the Legal System Sample*

**Table A3.** Legal Practitioners' Declared Country of Origin.

| Country of Origin | Frequency | (%) |
|---|---|---|
| Brazil | 15 | 6.8 |
| Canada | 1 | .5 |
| Denmark | 1 | .5 |
| Finland | 77 | 34.8 |
| France | 1 | .5 |
| Germany | 50 | 22.6 |
| Greece | 1 | .5 |
| Iceland | 1 | .5 |
| Ireland | 3 | 1.4 |
| Spain | 8 | 3.6 |
| Sweden | 51 | 23.1 |
| United Kingdom | 9 | 4.1 |
| United States | 3 | 1.4 |
| **Total** | **221** | **100** |

**Table A4.** Legal Practitioners' Declared Profession.

| Profession | Frequency | (%) |
|---|---|---|
| Barrister | 1 | .5 |
| Criminal Investigator | 1 | .5 |
| Forensic Investigator | 1 | .5 |
| Judge | 26 | 11.8 |
| Lawyer | 118 | 53.4 |
| Legal Assistant | 2 | .9 |
| Legal Investigator | 1 | .5 |
| Paralegal | 2 | 1 |
| Police and Crime Commissioner | 1 | .5 |
| Police Chief | 2 | .9 |
| Police Commissioner | 4 | 1.8 |
| Police Investigator | 6 | 2.7 |
| Police Officer | 27 | 12.2 |
| Prosecutor | 20 | 9 |
| Social Worker | 9 | 4.1 |
| **Total** | **221** | **100** |

# Appendix B

**Table B1.** Final Content Analysis Categories for Participants in the Child Current Sexual Abuse Group.

| | **Categories** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| Participants' answers coded per category | 2 | 20 | 17 | 14 | 219 | 13 | 10 | 76 | 87 | 306 | - |
| | 32 | 23 | 134 | 26 | - | 46 | 59 | 111 | 153 | - | - |
| | 101 | 45 | 153 | 200 | - | 51 | 67 | 151 | 170 | - | - |
| | 110 | 51 | 195 | 281 | - | 52 | 143 | - | - | - | - |
| | 113 | 54 | 214 | 318 | - | 77 | 148 | - | | - | - |
| | 126 | 58 | 324 | - | - | 137 | 166 | - | - | - | - |
| | 135 | 77 | - | - | - | 198 | 177 | - | - | - | - |
| | 137 | 87 | - | - | - | 261 | 204 | - | - | - | - |
| | 141 | 128 | - | - | - | - | 223 | - | - | - | - |
| | 153 | 137 | - | - | - | - | 224 | - | - | - | - |
| | 188 | 153 | - | - | - | - | 227 | - | - | - | - |
| | 195 | 160 | - | - | - | - | 249 | - | - | - | - |
| | 236 | 186 | - | - | - | - | 252 | - | - | - | - |
| | 249 | 188 | - | - | - | - | 258 | - | - | - | - |
| | 253 | 214 | - | - | - | - | 277 | - | - | - | - |
| | 256 | 236 | - | - | - | - | 286 | - | - | - | - |
| | 264 | 281 | - | - | - | - | 303 | - | - | - | - |
| | 269 | 309 | - | - | - | - | 319 | - | - | - | - |
| | 283 | - | - | - | - | - | 320 | - | - | - | - |
| | - | - | - | - | - | - | 235 | - | - | - | - |
| **Total answers per category** | **19** | **18** | **6** | **5** | **1** | **8** | **20** | **3** | **3** | **1** | **-** |

Note: it is possible for a participant's ID to be under more than one column, since one response could fit in more than one category.

**Table B2.** Final Content Analysis Categories for Participants in the Adult Current Sexual Abuse Group.

| | **Categories** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| Participants' answers coded per category | 1 | 30 | 1 | 33 | 98 | - | 16 | 25 | 220 | 48 | 3 |
| | 9 | 47 | 3 | 61 | 263 | - | 19 | 95 | - | 199 | 7 |
| | 22 | 48 | 40 | 119 | - | - | 35 | 206 | - | - | 38 |
| | 28 | 96 | 47 | 127 | - | - | 43 | 225 | - | - | 40 |
| | 41 | 104 | 48 | 139 | - | - | 57 | 244 | - | - | 48 |
| | 96 | 112 | 61 | 144 | - | - | 84 | 307 | - | - | 60 |
| | 144 | 202 | 64 | 171 | - | - | 109 | - | - | - | 116 |
| | 152 | 271 | 104 | 196 | - | - | 150 | - | - | - | 144 |
| | 168 | 307 | 139 | 213 | - | - | 157 | - | - | - | - |
| | 176 | - | 144 | 248 | - | - | 196 | - | - | - | - |
| | 187 | - | 147 | 251 | - | - | 222 | - | - | - | - |
| | 197 | - | 171 | 260 | - | - | 255 | - | - | - | - |
| | 229 | - | 184 | 285 | - | - | 262 | - | - | - | - |
| | 234 | - | 217 | - | - | - | 300 | - | - | - | - |
| | 282 | - | - | - | - | - | 307 | - | - | - | - |
| | 289 | - | - | - | - | - | 308 | - | - | - | - |
| | 304 | - | - | - | - | - | 315 | - | - | - | - |
| | 307 | - | - | - | - | - | 322 | - | - | - | - |
| | - | - | - | - | - | - | 323 | - | - | - | - |
| **Total answers per category** | **18** | **9** | **14** | **13** | **2** | **-** | **19** | **6** | **1** | **2** | **8** |

Note: it is possible for a participant's ID to be under more than one column, since one response could fit in more than one category.

**Table B3.** Final Content Analysis Categories for Participants in the Adult Historical Sexual Abuse Group.

| | **Categories** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| Participants' answers coded per category | 12 | 11 | 21 | 50 | 216 | - | 36 | 15 | 272 | 11 | 49 |
| | 29 | 63 | 24 | 63 | 296 | - | 42 | 56 | 291 | - | 201 |
| | 39 | 221 | 31 | 69 | - | - | 44 | 102 | 314 | - | 317 |
| | 66 | 310 | 50 | 82 | - | - | 62 | 226 | - | - | - |
| | 75 | 331 | 63 | 149 | - | - | 80 | 231 | - | - | - |
| | 82 | - | 85 | 165 | - | - | 88 | - | - | - | - |
| | 108 | - | 97 | 302 | - | - | 117 | - | - | - | - |
| | 114 | - | 138 | 310 | - | - | 140 | - | - | - | - |
| | 122 | - | - | 321 | - | - | 145 | - | - | - | - |
| | 125 | - | - | - | - | - | 146 | - | - | - | - |
| | 142 | - | - | - | - | - | 155 | - | - | - | - |
| | 154 | - | - | - | - | - | 161 | - | - | - | - |
| | 175 | - | - | - | - | - | 174 | - | - | - | - |
| | 191 | - | - | - | - | - | 183 | - | - | - | - |
| | 192 | - | - | - | - | - | 185 | - | - | - | - |
| | 212 | - | - | - | - | - | 203 | - | - | - | - |
| | 216 | - | - | - | - | - | 233 | - | - | - | - |
| | 254 | - | - | - | - | - | 243 | - | - | - | - |
| | 272 | - | - | - | - | - | 246 | - | - | - | - |
| | 284 | - | - | - | - | - | 247 | - | - | - | - |
| | 288 | - | - | - | - | - | 259 | - | - | - | - |
| | 305 | - | - | - | - | - | 279 | - | - | - | - |
| | 314 | - | - | - | - | - | - | - | - | - | - |
| | 317 | - | - | - | - | - | - | - | - | - | - |
| **Total answers per category** | **24** | **5** | **8** | **9** | **2** | **-** | **22** | **5** | **3** | **1** | **3** |

Note: it is possible for a participant's ID to be under more than one column, since one response could fit in more than one category.

# Chapter V

## Implanting False Autobiographical Memories of Repeated Events

Calado, B., Luke, T. J., Connolly, D. A., Landström, S., & Otgaar, H.

*This chapter corresponds to the following paper:*
Implanting false autobiographical memories for repeated events.
Memory. doi: 10.1080/09658211.2021.1981944

**Abstract**

Research to date has exclusively focused on the implantation of false memories of single events. The current experiment is the first proof of concept that false memories of repeated autobiographical experiences can be implanted using an adapted false memory implantation paradigm. We predicted that false memory implantation approaches for repeated events would generate fewer false memories compared with the classic implantation method for single events. We assigned English speaking participants to one of three implantation groups in our study: Standard, Repeated, and Gradual. A hundred and twenty participants underwent three interview sessions, with a 1 week interval between sessions. In the Standard group, we exposed participants to a single-event implantation method in all three interviews. In the Repeated group, the participants underwent a repeated event implantation method in all three interviews. The Gradual group also consisted of a repeated event implantation method; however, in the first interview alone, we suggested to participants that they had experienced the false narrative once. Surprisingly, within our sample, false memories rates in the Standard group were not higher compared with those in the Repeated and Gradual groups. Although sometimes debated, our results imply that false memories of repeated events can be implanted under lab conditions, likely with the same ease as false memories of single events.

***Keywords:*** false memory implantation, repeated events, adults' memory

## Opening

For over 25 years, researchers have examined how people can form rich false episodic memories in suggestive therapeutic or investigative interview contexts (Loftus & Pickrell, 1995; Scoboria et al., 2017). Research on false memory formation for autobiographical events is relevant for cases in which suggestive psychotherapeutic practices may have fostered false memories of child sexual abuse (see Maran, 2010; Lindsay & Read, 1994; Scoboria et al., 2017). False memories are probably responsible for only a small proportion of sexual abuse cases in the courtroom, but they are highly detrimental to the justice system. These cases can waste substantial public resources in costly and prolonged forensic investigations, leading to false accusations and potential miscarriages of justice (Loftus, 2005; see: Maran, 2010).

Some critics have aptly noted that victims of sexual abuse often experience and recall multiple episodes of abuse, rather than a single episode (Brewin & Andrews, 2017; Blizard & Shaw, 2019). However, to date, no experimental studies have demonstrated the implantation of false autobiographical memories of repeated events; thus, the existing literature may be limited in this important aspect. Here, we present the first study attempting to implant false autobiographical memories of repeated events.

## False Memory Implantation

To investigate whether false memories of entire autobiographical events could be implanted experimentally, Loftus and Pickrell (1995) designed the now-classic 'lost-in-the-mall' paradigm. In their study, the experimenter collected narratives of episodes from the participants' family that happened during the participants' childhood. The participants were then asked to report everything they could remember about these events. Unbeknownst to them, one event never happened (i.e. being lost in a shopping mall or big department store) and was actually manufactured by the researchers. In studies using this paradigm, a non-trivial percentage of participants claimed to have experienced this false event (on average, around 30%; Scoboria et al., 2017). However, as critics have pointed out, these studies only focused on the implantation of single events (see Table A1 for an overview of false memory implantation studies using adult participants).

Critics such as Blizard and Shaw (2019) have noted that false memory researchers have not 'been able to implant memories for repeated events, as is often the case with reported childhood sexual abuse' (p. 15). Similarly, Brewin and Andrews (2017) stressed that, 'A challenge for the future will be to demonstrate that it is possible to implant memories of a repeated event' (p. 20). Clearly, an experimental demonstration of the implantation of autobiographical false memories of repeated events is necessary in order

to address these limitations in the existing literature.

Indeed, implanting false memories of repeated events may be more difficult than implanting false memories of single events. Metacognitive processes might inhibit the formation of false memories of repeated events. That is, people may doubt that a suggested episode could have happened to them repeatedly without them having any script (as per schema theory; Farrar & Goodman, 1992) or gist trace (as per fuzzy trace theory; Brainerd & Reyna, 2012) about the event. Consequently, people may reject the suggestion of having experienced a repeated false event. Associative activation theory posits that, when a certain experience repeats itself, a script (i.e. the cluster of typical activities that occur during an event) of this experience is created containing details that are strongly interconnected with each other (Howe et al., 2009; Otgaar et al., 2019). In addition, having a script concerning a suggested event has been shown to facilitate the formation of implanted false memories (Otgaar et al., 2010; Pezdek et al., 2006).

## The Current Experiment

To directly address criticisms of the literature in the current experiment, we adapted the lost-in-the-mall paradigm and examined (1) whether rich autobiographical false memories of repeated events could be implanted; (2) whether the implantation of false memories of repeated events would be easier or more difficult than the implantation of single events; and (3) whether familiarisation with the event, prior to the suggestion that it occurred repeatedly, would facilitate the implantation of false memories of repeated events.

We examined two approaches to implant false memories of repeated events. Two of our experimental groups received false suggestions that a childhood event had happened repeatedly to them. In one group, in all three interview sessions, the participants received the suggestion that the false event had happened to them repeatedly. Almost the same procedure was followed in the other group, except that the participants in the first session were told that the event had happened once; then, in the second and third sessions, it was suggested to the participants that the event had happened to them repeatedly.

In the latter group, our aim was to examine whether the implantation of repeated events could be facilitated if we first established some familiarity with the false narrative before increasing the frequency of the event occurrences. As suggested by Mazzoni and Kirsch's (2002) metacognitive model, establishing such familiarity might contribute to the formation of a script of the false event, which would then promote false memory production (see also Ost et al., 2008). Furthermore, back to potential cases of false memory of repeated events in the 1980s and 1990s in the US, psychotherapy patients were fist familiarised with the possibility of being sexually abused as child. We predicted that it would be more difficult to implant false memories for repeated events than for

single ones. In addition, we expected that familiarisation with the false narrative would facilitate the implantation of false memories of repeated events.

## Method

We preregistered the methods and materials for the present study on the Open Science Framework (OSF). Before the data collection was complete and prior to the data analysis, we updated the registration to offer more specified hypotheses and corresponding statistical tests, since our initial registration was not sufficiently detailed about the statistical approach that would be used to test the predictions. Two additional registrations were made to update the randomisation procedures (i.e. to rebalance the groups after data exclusions) and to update the data coding procedures. These updates were created prior to the data analysis. All of this information and supplementary materials are available at https://osf.io/4fzht/?view_only=cf0035032a124786b7ccce67ec60e43c.

### *Participants*

We determined in advance that we would recruit participants until we reached a target sample size of $N$ = 120 participants with usable data. We based this target on what we considered to be feasible, given the available time and resources. Furthermore, to the best of our knowledge, the present study uses the largest sample in the literature on false memory implantation (see Table A1). With three measurements per participant ($N$ = 360 observations), a sensitivity analysis suggested that this sample size provides a power of .80 to detect the effects of false memory formation $f^2$ = .03 ($d$ = .32), assuming a conventional alpha level of .05. For this calculation, we used the pwr package (Champely, 2018) for R (R Core Team, 2018).

We recruited a total of 245 English speaking students from one of the universities associated with this project and from a music conservatory in the Netherlands. Our participants came from a variety of countries and ranged in age from 18 to 36 years (see Appendix A for more demographic information about our participants). Participants were excluded if they were found to (1) have actually experienced our false narrative (this was checked with the participants' parents); (2) be outside of the age range of 18–36 years old (this was assessed via our pre-screening demographic questionnaire); (3) be currently enrolled in the third year of a psychology bachelor's, master's, or doctorate programme at University X (also assessed via the pre-screening demographic questionnaire); (4) have spoken to family and friends about the narratives used in the study during the testing phase (we asked the participants in the debriefing phase of the study); or (5) have participated in other false memory implantation studies at University X (this was checked via email with the participants in the pre-screening phase). In addition, some parents did not provide childhood stories for some of our participants, which invalidated the continuation of their participation in the study. In addition, some participants did

not respond to our emails to schedule their interview sessions and consequently did not participate in the study.

In total, 125 participants were excluded. With the exception of one participant, these exclusions took place in the pre-screening phase. The exception was a participant who told us during the testing phase that his housemate had also participated in our study and had told him the real intent of the project. These exclusion criteria were in place to ensure that (1) the suggested narrative was indeed a false narrative for each participant; (2) the suggested narrative would have taken place a reasonably similar number of years ago for each participant; (3) the participants were less likely to have received relevant education about false memory phenomena; (4) the participants did not discover that the suggested narrative was false during the testing period; and (5) the participants had not been exposed to similar procedures. After exclusions, the final sample size was $N$ = 120 participants. The participants were rewarded for their participation at the end of the final interview session with a voucher of €15 or academic credits. This study was approved by the standing ethical committee of University X.

***Design***

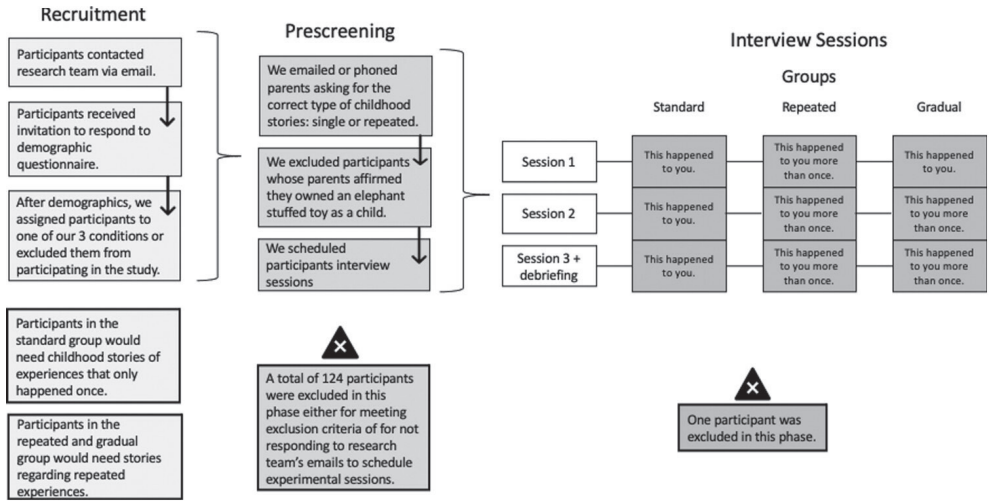
We used a three (Implantation method: Standard, Repeated, and Gradual; between subjects) by three (Time: Interview 1, Interview 2, Interview 3; within subjects) mixed design. The participants were randomly assigned to one of the three different groups (Standard, $N$ = 40; Repeated, $N$ = 41; Gradual, $N$ = 39). The participants in the Standard and Repeated groups consistently heard from the interviewer, in all three interview sessions, that they had experienced the false narrative (as well as their true experiences) in a given fixed frequency, either once or more than once. However, the participants in the Gradual group heard in the first interview session that the false narrative (and true experiences) had occurred to them once, while in the second and third interview sessions, the interviewer suggested that all of the childhood stories (including the false narrative) had actually happened more than once.

***Procedure***

Figure 1 provides a diagram of the procedures. We advertised our study as 'Childhood Memories', stating that we were interested in learning how well people could remember their childhood experiences. Students who were interested in participating in our study were asked to contact us via email and received a link to an online demographic questionnaire. After we received their demographic information, we randomly assigned the participants to one of the three experimental groups. Randomisation was necessary at this point because the screening procedure was different based on the participant's assigned group. More specifically, we contacted the participants' families to collect either stories of childhood events that happened once (for the Standard group) or stories of childhood events that had happened several times (for the Repeated and Gradual groups).

In order to collect true stories from the participants' childhood, as well as to check whether our false narrative had in fact happened to them, we contacted the family member that had been nominated by the participants in their demographic questionnaire, by email or phone. In the majority of cases, the participants indicated their parents (97.5%); only one participant had indicated a brother and mother, and two participants had indicated their sisters.

**Figure 1.** Experimental Procedure.




The eight interviewers in this study were students from the Legal and Forensic Psychology Master's programmes at University X, who worked on our project for their master's thesis. For this reason, they were aware of all the hypotheses and the design of the experiment. They were trained for a total of 40 hr spread over a period of 2 weeks to become familiar with the interview script that had to be followed strictly in order to avoid excessive differences between the interviewers. We trained the interviewers so that they could conduct the interviews for all three experimental groups.

All the participants underwent three interview sessions that were audio recorded, with a 1-week interval between each session. The debriefing phase took place at the end of the third session. In the interview sessions, the experimenter told the participants that their parents had reported to us that all the narratives had taken place in the participants' lives between the ages of 5 and 10 years. For all three interviews, the childhood stories were told in the following order: (1) true story; (2) true story; (3) false story; and (4) true story. More specifically, in the false narrative that (supposedly) had only happened once, we told the participants the following:

> So, your mum told us that, when you were a child, you lost one of your favourite cuddling toys. It was an elephant that you really liked. According to your mum, it was very hard for you to lose this toy, you got really sad, and cried a lot. Tell me what you remember about this event.

In the narratives for the repeated (false) childhood event, we told the participants the following:

> So, your mum told us that, when you were a child, it happened to you more than once that you lost one of your favourite cuddling toys. It was an elephant that you really liked. According to your mum, it was very hard for you to lose this toy, you got really sad and, cried a lot. This happened to you more than once. Tell me what you remember the most about the times this happened.

In this particular example, we refer to the participant's mother; however, this was adapted according to the relative the participants had referred to us.

***Interview 1***

The interview setting was always a quiet, simple room with two chairs and one table. After the participants were welcomed to the interview setting, the interviewer explained to them the supposed aim of the study, using our cover story. The interviewer reminded the participants that all three interview sessions would be voice recorded for data collection purposes; in addition, the interviewer handed to the participants the study information letter and the informed consent form. Subsequently, the interviewer told the participants that the research team had contacted their families in order to collect childhood episodes that had happened to them.

For those in the Standard group, the interviewer narrated the events, while explaining to the participants that each event had 'happened to them' and asking them to describe everything they could remember about that occurrence. The wording used by the interviewer specifically implied that the events had only happened once. In the Repeated group, the participants were informed that the family member who provided the narratives had attested that all of the narrated events had happened more than once. In the Gradual group, in the first interview, the participants similarly learned that these events had 'happened to them', just as in the Standard group.

In line with past studies (displayed in Table A1), whenever the participants had any difficulty with retrieving memories of any of the episodes, the interviewer encouraged them to try to remember again. Moreover, the interviewer told the participants that it is common to forget certain events that happened a long time ago. After this encouragement, if the participants still could not remember the given experience, the interviewer used

context reinstatement (i.e. mentally re-establishing the environment and emotional state) and guided imagery techniques (Goff & Roediger, 1998; Lindsay & Read, 1994), as previous studies have done. At the end of the first interview, the interviewer encouraged the participants to think about the events they had been presented with. The interviewer also asked them not to talk to anyone about the experiment or the stories they were trying to recall.

***Interview 2***

In the second interview, the interviewers narrated verbatim the same four stories to the participants that they had heard in Interview 1. For both the Standard and Repeated groups, the frequency of the supposed events was the same as in Interview 1: single events for the Standard group and repeated events for the Repeated group. Unlike these groups, in the second interview, the Gradual group's frequency of events changed in comparison with the first interview for the same group. That is, the interviewer said the following to the participants in the introduction of the second interview session:

> Last week, I told you about four events that your mother[8] told us you experienced a few times when you were around the age of 5 to 10 years old. It turns out that there's a really high chance that they happened to you more than once. As I said last week, I am particularly interested in what you can remember of those events compared with your mother's report. Hopefully, you could remember more details throughout the week. So, I'm going to read you those events all over again and, after reading them, I ask you to tell me everything you can remember. Do you understand? Ok then, we're going to start now.

In all the groups, the interviewers applied the same procedure of context reinstatement and imaginary inflation that had been used in Interview 1, when necessary. At the end of the session, the interviewer encouraged the participants to think about all the stories that had been presented to them. The interviewer also asked the participants not to talk to anyone about the experiment and to think about the experiences they had been unable to recall.

***Interview 3***

For all three groups, the third interview followed the same steps as the second interview. At the end of this session, the participants were debriefed regarding the true nature of the study. The interviewers asked the participants if they could guess which of the four stories, they had heard in the experiment, was actually false.

8 We used 'mother' as an example, but this could have been any family member indicated by the participants as their contact person.

## Materials

### *True Narratives*

We collected three truly experienced stories from the participants' families. More specifically, for the participants assigned to the Repeated and Gradual groups, we collected recurrent childhood experiences; for the participants placed in the Standard group, we gathered childhood events that had only happened once. All of the collected childhood events should have happened to the participants between the ages of 5 and 10 years in order to keep the events consistent with the same time frame we had chosen for our false narrative.

### *False Narrative*

To find a suitably plausible negative experience to falsely suggest in our experiment, we conducted a pilot study ($N$ = 20) with co-workers from University X to examine 10 different possible events to be used in our study (https://osf.io/4fzht/?view_only=cf0035032a124786b7ccce67ec60e43c). These events have been used in the literature in previous research (see Table A1). We asked the participants to score each of the 10 different narratives from 1 to 7 in terms of plausibility (1 = not plausible, 7 = extremely plausible), and emotional valence (1 = positive, 7 = negative). The events given to the participants to assess included events such as being lost in a shopping mall and going on a hot air balloon ride, which have been used in previous research (e.g. Loftus & Pickrell, 1995; Wade et al., 2002). For ethical reasons, we aimed to choose a childhood experience that was rated as plausible or extremely plausible (scoring from 5–7) and rated as negative, but not extremely negative (scoring from 5–6), which could be adapted for both the Repeated and Standard groups.

Pilot participants rated 'losing a toy' as the most plausible event ($M$ = 6.85, $SD$ = 0.36), combined with a moderately negative valence ($M$ = 5.37, $SD$ = 1.38). Based on this result, we created the following false narrative: Between the ages of 5 and 10 years, the participant had lost an elephant cuddly toy (i.e. a stuffed animal), was very upset about it, and cried a lot. For situations in which the participants did not indicate a location where the event might have occurred, we created three possible scenarios in which this event could have happened, for the interviewers to suggest. To be specific, the interviewers suggested that the participants could have lost their elephant in a park, at school or in a supermarket. The participants in the Repeated and Gradual groups were told that they had lost the elephant several times, each time recovering the toy, until they lost it for good.

### *Autobiographical Belief and Recollection Ratings*

After the presentation of a true or false narrative, the participants received a questionnaire, in which they could rate their own recollection of and belief in the events on a scale from 1 ('I do not recall this event' or 'This did not happen to me') to 8 ('I remember this event

completely' or 'I am completely confident that this happened to me'; see Otgaar et al., 2013; Scoboria et al., 2004). For each narrative, the participants rated their recollection and belief separately. The interviewer explained to the participants that for 'recollection', they should consider any mental images that occurred to them concerning each narrative; for 'belief', the interviewer explained that we wanted to know how confident they were that they had experienced the narratives.

***Researcher Coding***

In addition to the participants' self-scoring, we used the approach reported by Scoboria and colleagues (2017) to code the content of the false narratives. This coding takes under consideration both belief in and recollection of false memories. We adapted the rating form for the participants in the Repeated and Gradual groups by adding extra questions and duplicating some of the original questions from the work by Scoboria et al. (2017). We elaborate on the coding procedures in the following subsections. The complete research rating instrument is available at https://osf.io/4fzht/?view_only=cf0035032a124786b7ccce67ec60e43c.

***Memory and Belief Quality in the Repeated and Gradual Groups***

In the Repeated and Gradual groups, participants (self-rating) who had up to three occurrences of the false narrative to report ranked each occurrence from most to least vivid and believed. In the self-report, the first scored occurrence of the false event corresponded to their first most vivid and believed occurrence. They then rated their second most vivid and believed occurrence, and finally their third most vivid and believed occurrence. The researcher also followed this system when assessing the participants' memory statements and coding the research coding instrument.

***Hypotheses***

We had three hypotheses:

(1) Acceptance and false memory rates for the Standard group were expected to be similar to the rates presented in Scoboria and colleagues' (2017) mega-analysis for the same dependent variables (Acceptance = 69.7% and False Memory = 30%). Using a chi square test, we compared our results for the rates of acceptance and false memories with those obtained in the general literature, to ensure that our results in the Standard group were comparable to past false memory implantation attempts.
(2) We predicted that inducing the participants to believe and recall the false narrative in a repeated-occasion format would be more difficult, compared with a single-occasion format. That is, we expected that both the participants' self-reports and the researcher ratings would show that belief in and recollection of the false narrative would be lower in the Repeated and Gradual groups than in the Standard group of this experiment.

To test this prediction, we used linear mixed-effect models to compare the data for belief in and recollection of the false narrative from the Repeated and Gradual groups with that from the Standard group across the three interviews. We used this statistical approach both for the participants' self-reports and for the researcher ratings.

(3) We predicted that the gradual introduction of the suggestion of a false narrative would facilitate a more effective implantation compared with repeatedly suggesting that the false event had occurred from the outset. Therefore, the belief and recollection from the Gradual group would be higher than those from the Repeated group. In the Gradual group, we attempt to create familiarity with the false narrative before introducing the frequency with which it (supposedly) had happened. The linear mixed-effect models described above also provided a comparison of false beliefs and false memory rates for the Repeated and Gradual groups.

## Coding and Analysis

### *Researcher Rating*

The eight researchers who interviewed the participants coded their data after every interview while following the same coding system, as in the previous mega-analysis of false memory implantation studies by Scoboria et al. (2017). As shown in Table 1, the coding used by Scoboria et al. shows only one variable that purely measures the rates of belief, but shows three measures for recollection. In addition to coding the variables displayed in Table 1, we coded several exploratory variables (not reported here, in the interest of space), similar to those measured by Scoboria et al. (2017) (e.g. the emotional content of the narrative).

**Table 1.** Variables Used to Measure False Memory Rates by Scoboria et al. (2017).

| Item | Variables | Scale |
|---|---|---|
| 1. Acceptance | 1. Participant's acceptance of the specific details of the false event that were provided to her/him | 0. Outright rejection;<br>1. Minimal acceptance/accept parts, reject parts;<br>2. Moderate acceptance/acceptance with no active rejection;<br>3. Complete acceptance. |
| 2. Elaboration | 2. Information given by participant beyond what was provided to her/him | No/Yes |
| | 3. Amount of elaboration | 1. Minimal;<br>2. Moderate;<br>3. Substantial. |
| 3. Imagery | 5. Sensory imagery generated by participant | No/Yes |
| | 6. Amount of sensory imagery generated by participant | 1. Small amount of imagery (about 1 image);<br>2. Moderate amount of imagery (about 2–4 images);<br>3. High amount of imagery (about 5 or more images) |
| 4. Event rejection | 7. Whether participant verbally states that she/he did not have a memory of the false narrative | No/Yes |

Because we were interested in repeated events, we measured our variables at the level of each occurrence of the false event (i.e. up to three occurrences) within each interview (i.e. three interviews). Simple direct statements such as 'I remember this happened to me' were not sufficient for us to conclude that a participant had formed false memories of the false narrative. Rather, when scoring the data, we examined their statements in full, as other elements might indicate whether a false memory was formed, such as elaborating on details (e.g. 'I remember it was raining that day.' or 'The park was full of children when we were there.'). We used the same false memory categories as Scoboria et al. (2017) to classify the participants' coded statements. The categories were as follows: Robust False Memory, Full False Memory, Partial False Memory, Accepted False Memory, No False Memory, and Rejected False Memory. Table A2 provides descriptions of each false memory category.

For a participant's memory report to qualify as a Robust False Memory, a high level of acceptance had to be present, along with a moderate level of elaboration and imagery. To be classified as a Full False Memory, the results had to meet a moderate level of acceptance, elaboration and imagery. For the Partial False Memory category, the results needed to display a moderate level of acceptance with any level of elaboration and imagery lower than a moderate level. To qualify as an Accepted False Memory, the results had to solely show a moderate level of acceptance and an absence of elaboration or imagery. To be coded as No False Memory, the results had to show any level of acceptance below a moderate level and an absence of elaboration or imagery. Finally, a participant's result qualified as a Rejected False Memory when the results showed no level of acceptance and an absence of elaboration and imagery.

***Interrater Reliability***

To perform our interrater reliability coding, we trained two professional research assistants who were blind to the hypothesis of the study for 20 hours. We instructed them to code all the interview sessions in the same fashion used by the eight interviewers. The two raters studied the coding system used in Scoboria et al. (2017); they also had access to both the transcripts and the audio of the interview sessions in order to fill in the 'researcher coding forms'.

We used the data corresponding to 42 (35.8%) of the participants[9], selected randomly, to calculate the interrater reliability for the researcher's rating. As the index of agreement between the two raters for the false memory category coding, we used Gwet's AC, with linear weights for the ordinal categories (Gwet, 2008). Gwet's AC results can be interpreted in a manner similar to Cohen's kappa. The levels of interrater reliability displayed in Table 2 indicate that the agreement between the two raters' coding varied from moderate to high agreement (Gwet, 2008).

9 Out of 118 participants whose researcher rating data was available to us, sound files corresponding to two participants were corrupted and could not be accessed by the interrater reliability coders. Since the interviewers had access to both the transcript and the audio of the participants to code their memory reports, we decided to restrict the interrater reliability analysis to 118 participants instead of 120.

**Table 2.** Interrater Reliability for Researcher Coding.

| Variable | Gwet's AC | 95% CI |
|---|---|---|
| Participant's acceptance of the specific details of the false event that were provided to her/him | 0.52 | 0.44, 0.59 |
| Information given by participant beyond what was provided to her/him | 0.71 | 0.63, 0.80 |
| Amount of elaboration | 0.57 | 0.33, 0.81 |
| Sensory imagery generated by participant | 0.63 | 0.49, 0.76 |
| Amount of sensory imagery generated by participant | 0.64 | 0.37, 0.67 |
| Whether participant verbally stated that she/he did not have a memory of the false narrative | 0.78 | 0.72, 0.85 |

## Results

### *Summary*

Given the quantity and depth of the data produced by this study, we will begin with a brief narrative summary of the results, followed by formal statistical analyses. To provide a straightforward summary, we will assess the formation of false memories using the coded statements the participants provided in the interviews. In the Standard group, 25.64% (10/39) of the participants provided a statement in at least one of the three interviews that was classified as a Partial, Full or Robust False Memory. Thus, it appears that a non-trivial minority of the participants formed false memories of the suggested event. However, contrary to our expectations (Hypothesis 1), this rate was lower than the overall rate found in the literature (see below; Scoboria et al., 2017).

In the Repeated group, 34.14% (14/41) of the participants formed false memories (by this operationalisation); of these, 14.28% (2/14) provided more than one such statement within a single interview, indicating that they had formed false memories of repeated events. In a similar pattern, 36.84% (14/38) of the participants in the Gradual group formed false memories; of these, 28.57% (4/14) formed false memories of repeated events. Although the rate of false memory formation for repeated events was relatively low, having the interviewers suggest that the target event occurred more than once did not decrease the overall rate of false memory formation, contrary to our expectations (Hypotheses 2 and 3). These results are corroborated by the participants' self-reports (see below).

In the Standard group, 76.92% (30/39) of the participants provided at least one statement indicating acceptance of a false memory of the suggested event. By this measure, the participants in the Repeated (75.61%, 31/41) and Gradual (81.57%, 31/38) groups accepted at least one occasion of the suggested event at similar rates. A substantial minority of the participants in the Repeated (24.39%, 10/41) and Gradual (36.84%, 14/38) groups provided more than one statement within a single interview that indicated at least acceptance of the event, which suggests that false beliefs had been formed for repeated events (see Table C3).

In short, suggesting that the false event happened repeatedly (either from the outset of the interviews or gradually) led to some of the participants believing and/or remembering that they had experienced the event several times, without reducing the overall rate of false memory formation as we had expected. Thus, this study demonstrates the feasibility of implanting false memories of repeated events.

***Overview of Statistical Results***

Three dependent variables provided information about the participants' formation of false memories: (1) self-reported recollection of the false event, (2) self-reported belief that the false event occurred, and (3) researcher coding of participants' recorded narratives for indications of false memories. Furthermore, in the Repeated and Gradual groups, we referred to the quality of false recollections in terms of 'vividness'. The participants in these groups rated up to three different occasions of the false narrative event in order from most to least vivid.

Each measure of false memory formation is subject to different limitations and sources of error. Some participants tended to be less verbally descriptive about their childhood narratives over time (possibly following an implicit norm of avoiding redundancy). That is, a participant's self-reports about the vividness of her or his (false) memories could remain somewhat consistent or could increase over time; however, researcher coding might suggest that this participant, for example, had formed partial false memories in the first interview and had accepted false memories in the second and third interview. Thus, examining all three relevant measures and interpreting their results taken together leads to the most trustworthy conclusions.

Here, we start with a subsection reporting the results corresponding to Hypothesis 1, which refers to the false memory rates of the Standard group – as coded by the researcher – being similar to the literature benchmark rates (Scoboria et al., 2017). Subsequently, we will present descriptive results of the data. For the self-reported data and the researcher coding, we will present inferential analyses corresponding to Hypothesis 2, which refers to false memories being more difficult to implant under the Repeated and Gradual groups' conditions compared with the Standard group's conditions; and Hypothesis 3, which involves comparing the Repeated and Gradual groups and predicting that false memory implantation will be more successful in the latter.

***Comparison with Past Implantation Studies***

To address Hypothesis 1, we examined the approach used for the Standard group in the present experiment, which is equivalent to the implantation approach of all the studies reported in Scoboria et al. (2017). For this comparison, we counted the number of participants who provided a report in any interview session that met the criteria of Partial, Full or Robust False Memory. In the Standard group ($n$ = 39) of this experiment, 25% (10) of the participants formed false memories and 77% accepted the false narrative

as true, as assessed by the researcher rating. These false recollection rates were lower than the literature benchmark of 46% for false memory ($N$ = 423; Scoboria et al., 2017), $\chi^2$ (1, $N$ = 462) = 5.92, $\phi$ = 0.12, p = .01. There was a small, nonsignificant difference between the literature benchmark of 69% and our experiment in terms of false belief rates ($N$ = 423; Scoboria et al., 2017), $\chi^2$ (1, $N$ = 462) = 1.11, $\phi$ = 0.04, $p$ = .30.

***Researcher Ratings***

Due to damaged audio files, the interviews for two participants (one in the Standard and one in the Gradual group) could not be coded. Thus, the group sizes available for coding were as follows: Standard $n$ = 39; Repeated $n$ = 41; and Gradual $n$ = 38.

Each participant in the Standard group ($n$ = 39) provided three measures of belief and three measures of recollection (one in each of the three interviews). Each participant in the Repeated group ($n$ = 41) provided nine measures of belief and nine of recollection (three incidents in each of the three interviews). Each participant in the Gradual group ($n$ = 41) provided seven measures of belief and seven of recollection (one incident in Interview 1 and three incidents in Interviews 2 and 3).

***Descriptive Analysis***

Table C2 provides the descriptive results of the researcher ratings for each participant measurement. The table displays the count for the cases pertaining to each category and, in parentheses, the percentages of the total number of cases for each implantation category. Figure C1 shows the development of false memory implantation over time (Interviews 1, 2, and 3). It also displays the rating attributed by the researcher to the participants' response on each occasion when the false narrative was suggested by the interviewer across the three interview sessions.

These results show that a non-trivial number of false memories fell into the 'Accepted False Memory' category by Interview 3 (Standard group: 38%; Repeated group: 22%; Gradual group: 26%). A smaller number of false memories fall into the 'Partial False Memory', 'Full False Memory', and 'Robust False Memory' categories. Moreover, 25% of the participants in the Repeated group and 34% of the participants in the Gradual group described two or more occurrences of the false narrative that were categorised under the researcher rating into either the 'Accepted False Memory' category or the 'Robust False Memory'.

Table C3 provides a detailed report for each participant who described two or more occurrences of the false events that the researcher classified as Accepted, Partial, Full or Robust false memory. There were different 'evolution' patterns of the false memories. As shown in Table C3, a few participants' statements suggested that their false memories were decreasing in vividness over time. For instance, the researcher assessed that, in the first interview, participant 2 in the Repeated group offered verbal information that matched a Robust false memory classification. However, in participant 2's second and

third interview, the researcher considered the information provided by the participant to fit into the Accepted false memory category, which is inferior to the Robust false memory classification in terms of false memory richness. Although it is not our objective to thoroughly attribute a justification for the manner in which the participants chose to communicate their memories, we do consider that it is possible that the participants avoided redundancy in their statements. For instance, participant 2 says in the third interview, 'I remember it. I think it is very blurry. Like I said last week. I can still remember mum talking to me about the lost elephant.'

### *Inferential Analysis*

**Analytic Approach**. For the participants in the Repeated and Gradual groups, the researcher rated up to three occurrences of the false narrative. The researchers also rated the quality of these occurrences (first most vivid, second most vivid, and third most vivid). The participants narrated their memories of the different occurrences of the false narrative while ordering them from the one they remembered the most to the one they remembered the least.

To assess the false memory formation elicited by the implantation approaches (Standard, Repeated, and Gradual), we used cumulative-link mixed-effects models with the ordinal package (Christensen, 2019) for R. For this model, we analysed the six false memory categories displayed in Table A2. We applied the same rationale to the researcher coding as we did to the participants' self-rating. Hence, the model for Hypothesis 2 included fixed effects for condition (Standard, Repeated, and Gradual) and interview session (1, 2, and 3). To test Hypothesis 2, we compared all three groups by comparing the first most vivid occurrence from the participants in the Repeated and Gradual groups with the only occurrence the participants reported in the Standard group. The model also included numeric factors (i.e. 0: Interview 1; 1: Interview 2; and 2: Interview 3) and random intercepts for the participants.

To test Hypothesis 3, we included in the model fixed effects for the Repeated and Gradual groups in all sessions (1, 2, and 3) and occurrences of the false narrative (most vivid, second most vivid, and third most vivid). With the exception of the first interview in the Gradual group, both the Repeated and Gradual groups allowed the participants to report their responses to up to three occurrences of the false narrative. For this reason, we included fixed effects in the model we used to assess Hypothesis 3 and compared both implantation methods. The fixed effects for the group and the number of occurrences used treatment contrasts (dummy coding), with the first listed level for each factor being taken as the reference group. This model also included numeric factors (i.e. 0: Interview 1; 1: Interview 2; and 2: Interview 3) and random intercepts for participants.

**Data Transformation.** The participants had up to three occurrences to rate their responses to the false narrative; consequently, the researcher also had up to three occurrences of the false narrative to rate. Whenever a participant did not describe a false occurrence, we treated that particular occurrence as 'no memory' (see Table A2). This was always the case for the second and third most remembered occurrences for the Standard group, considering that we did not intend to implant false memories of repeated events in this group. It was also the case for the second and third most vivid occurrence in the first interview session of the Gradual group. Other than this, every time a participant recalled, for instance, a first and second occurrence, but not a third, the third occurrence in the example was treated as a missing value and therefore as 'no memory'.

**Model Selection.** Similar to our model selection for the analysis of the data from the self-reported rating, we took a hierarchical approach fitting two models for Hypothesis 1 and Hypothesis 2, and a three-way interaction model for Hypothesis 3. The models consisted of one model with main effects as the fixed factors, a second model with two-way interactions for Hypotheses 2 and 3, and a third model with three-way interactions solely for Hypothesis 3.

When assessing Hypothesis 2, the second model that included two-way interactions fit the data better than the first model, $\chi^2$ (2) = 7.25, $p$ = .02.

When assessing Hypothesis 3, the second model, which included two-way interactions, fit the data better than the first model, $\chi^2$ (5) = 12.50, $p$ = .02. The third model did not fit the data better than the second model, $\chi^2$ (2) = 0.53, $p$ = .76. Thus, we retained the model that included the two-way interactions. We only report the results for the retained model here, but information about all the other models used in this study is available in the supplementary material (see https://osf.io/4fzht/?view_only=cf0035032a124786b7ccce67ec60e43c).

## Modelling Results

**Hypothesis 2.** This hypothesis questioned whether false memories of repeated events would be more difficult to implant than false memories of single events. Hence, in this analysis, we compared the results from the Standard group with the results from the first occurrence of the false narrative in the Repeated and the Gradual groups. Consistent with the participants' self-rating results, this model showed that false memory formation in the Repeated and Gradual groups was not significantly lower than in the Standard group. Furthermore, as displayed in Table 3, false memory formation appeared to increase significantly over time in the Repeated group, compared with the Standard group.

**Table 3.** Cumulative-Link Mixed-Effects Model Results Comparing Participants' False Memories in the Standard, Repeated and Gradual Groups.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***z*** | ***p* value** |
|---|---|---|---|---|
| Repeated group | -1.28 [-2.85, 0.29] | 0.80 | -1.58 | .11 |
| Gradual group | -0.02 [-1.63, 1.59] | 0.82 | -0.03 | .97 |
| Interview sessions | -0.03 [ -0.48, 0.42] | 0.23 | -0.12 | .89 |
| Repeated group × interview sessions | 0.88 [0.23, 1.53] | 0.33 | 2.60 | <.01 |
| Gradual group × interview sessions | 0.29 [-0.36, 0.94] | 0.33 | 0.86 | .38 |
| **Random effects** | ***SD*** | | | |
| Participants | 2.90 | | | |

**Hypothesis 3.** In Hypothesis 3, we examined whether the approach we employed in the Gradual group would elicit false memories in more participants compared with the method we used in the Repeated group. Table 4 shows that the participants in the Gradual group were slightly more likely to form at least one false memory in the first interview session. However, the tendency for the participant's false memories to increase over time is lower in the Gradual group compared with the Repeated group.

**Table 4.** Cumulative-Link Mixed-Effects Model Results Comparing Participants' False Memories in the Repeated and Gradual Groups.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***z*** | ***p* value** |
|---|---|---|---|---|
| Gradual group | 1.03 [0.05, 2.01] | 0.50 | 2.07 | .03 |
| Interview sessions | 0.81 [0.81, 0.81] | <.01 | 354.13 | <.0001 |
| Occurrence 2 | -1.35 [-1.35, -1.35] | <.01 | -375.33 | <.0001 |
| Occurrence 3 | -0.89 [-0.89, -0.89] | <.01 | -293.39 | <.0001 |
| Gradual condition × interview sessions | -0.54 [-0.91, -0.17] | 0.19 | -2.79 | .005 |
| Gradual condition × Occurrence 2 | 0.66 [0.01, 1.31] | 0.33 | 2.00 | .44 |
| Gradual condition × Occurrence 3 | 0.11 [-0.54, 0.76] | 0.33 | 0.34 | .72 |
| Interview sessions × Occurrence 2 | -0.68 [-0.68, -0.68] | <.01 | -199.41 | <.0001 |
| Interview sessions × Occurrence 3 | -0.83 [-0.83, -0.83] | <.01 | -274.44 | <.0001 |
| **Random effects** | ***SD*** | | | |
| Participants | 2.48 | | | |

## Participants' Self-Report Ratings

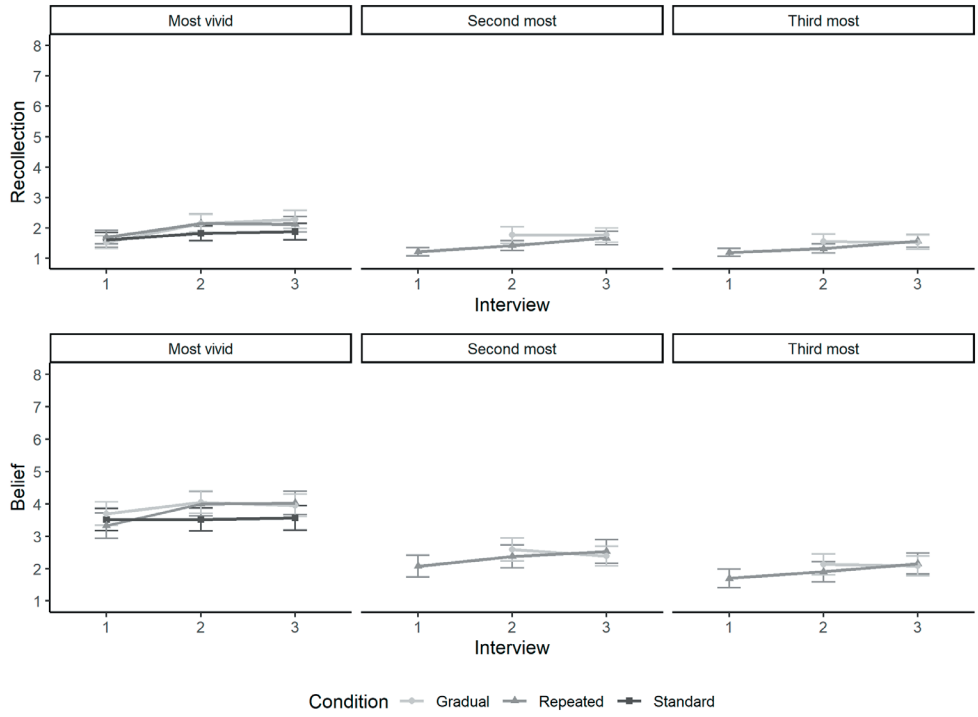
### *Descriptive Analysis*

The average of the participants' self-reports across all three conditions were higher for belief than for recollection. False recollection and false belief ratings for the second most and third most vivid memories were lower compared with the most vivid memory. If no one formed false memories or beliefs for the repeated events, the second and third most vivid occurrences would have means of 1.00 for recollection and belief, with no

variance. However, it is clear that false recollection and false belief self-reporting in the Repeated and Gradual groups are not constrained to the negative endpoint of the scale for the second and third most vivid occurrences. Thus, the participants in those conditions sometimes reported that they remembered or believed that the suggested false event occurred more than once.

Figure 2 shows that false recollection and false belief ratings tended to increase over time in the Gradual and Repeated groups. It is notable that, for both the false recollection and false belief measures, the mean values at all time points and in all conditions are fairly low on the respective scales, never rising above the midpoint.

**Figure 2.** Mean Self-Reported Recollection of and Belief in the False Narrative.




Note: Error bars represent standard errors.

### *Inferential Analysis*

**Analytic Approach.** We used linear mixed-effects modelling to assess the success of the implantation approaches (Standard, Repeated, and Gradual). We used the lme4 package (Bates et al., 2015) for R. This package was complemented with the lmerTest package (Kuznetsova et al., 2017) to obtain Satterthwaite approximated degrees of freedom. Our models used restricted maximum likelihood (REML) estimation. Each participant provided up to nine measures of belief and nine of recollection (described in more detail

below), providing up to $N = 1080$ valid observations for each measure. The model for Hypothesis 2 included fixed effects for condition (Standard, Repeated, and Gradual) and interview sessions (1, 2, and 3). Even though the participants in the Standard group only reported on one occurrence of the false narrative, such fixed effects allowed us to compare all three groups (Hypothesis 2) for the first most vivid occurrence in the Repeated and Gradual group, and for the only occurrence in the Standard group.

For Hypothesis 3, the model included fixed effects for condition (Repeated and Gradual), interview session (1, 2, and 3), and occurrences (most vivid, second most vivid, and third most vivid). We added the three occurrences of the false narrative only to test Hypothesis 3, because this hypothesis compares the Repeated and Gradual groups. With the exception of the first interview session for the Gradual group, the participants in the Repeated and Gradual groups were able to describe up to three occurrences of the false narrative. The fixed effect for the interview sessions used treatment contrasts (dummy coding) and numeric factors (i.e. 0: Interview 1; 1: Interview 2; and 2: Interview 3). The fixed effects for occurrences also used treatment contrasts (dummy coding), with the first listed level for each factor being taken as the reference group. The models included random intercepts for participants[10].

**Model Selection.** We took a hierarchical approach to model selection. In the first step, we fit a model that included only main effects for each of the fixed factors. In the second step, we added the two-way interactions. We compared each model with the previous step using a likelihood ratio test. These comparisons required refitting the models using the maximum likelihood instead of REML, but all of the reported results use the REML fittings.

When assessing Hypothesis 2, for the recollection measure, the second model did not offer a significant improvement compared with the first model, $\chi^2$ (2) = 2.21, $p$ = .33. The same held true for the belief measure: the second model (including two-way interactions) did not fit the data significantly better than the first model, $\chi^2$ (2) = 2.38, $p$ = .30. Thus, for both self-reported measures (recollection and belief), we retained the model with the main effects only. We report the results for those retained models here, but information about the other models is available in supplementary material (see https://osf.io/4fzht/?view_only=cf0035032a124786b7ccce67ec60e43c).

When assessing Hypothesis 3, for the recollection measure, neither the second model, $\chi^2$ (5) = 3.07, $p$ = .68, nor the third model (including three-way interactions), $\chi^2$ (2) = 2.81, $p$ = .24, fit the data better than the first model. For the belief measure, again, neither the second nor the third model fit the data significantly better than the

10 This is a different analysis than what we had proposed in our first pre-registration. We realised before the end of data collection that our original analytic strategy (i.e. mixed ANOVA) would not be as informative as a multilevel modelling approach. We therefore revised our analysis plans for both the self-reported data and the researcher coding. We registered these changes on OSF prior to data collection: https://osf.io/4fzht/?view_only=cf0035032a124786b7ccce67ec60e43c.

first model. For this reason, we report in this section the results of the first model, for both the recollection and belief measures.

### Modelling Results for Participants' Self-Report Ratings

**Hypothesis 2.** With this hypothesis, we aimed to assess whether false memories of repeated events were more difficult to implant than false memories of single events. Here, we compared the participants' self-rated recollection and belief ratings in the Standard group with those in the Repeated and Gradual groups, specifically for the participants' most vividly recalled event (excluding the second and third most vivid events). The results are presented in Tables 5 and 6. Against our predictions, neither the Repeated nor the Gradual group demonstrated significantly lower ratings of recollection or belief compared with the Standard group. Furthermore, in all three groups, the participants tended to rate their recollection of the false narrative higher in their second and third interview session than they did in their first interview session. There was a similar but nonsignificant tendency for belief ratings to increase over the interview sessions.

**Table 5.** Linear Mixed-Effects Model Results for False Recollection Ratings.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***t*** | ***df*** | ***p* value** |
|---|---|---|---|---|---|
| Intercept (Standard) | 1.53 [1.08, 1.98] | 0.23 | 6.55 | 137.91 | <.0001 |
| Repeated group | 0.21 [-0.40, 0.82] | 0.31 | 0.68 | 116.99 | .49 |
| Gradual group | 0.21 [-0.42, 0.84] | 0.32 | 0.68 | 116.99 | .49 |
| Interview sessions | 0.23 [0.11, 0.35] | 0.06 | 3.56 | 239.00 | <.001 |
| **Random effects** | ***SD*** | | | | |
| Participants | 2.84 | | | | |

Note: Coefficients are displayed with 95% CIs.

**Table 6.** Linear Mixed-Effect Model Results for False Belief.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***t*** | ***df*** | ***p* value** |
|---|---|---|---|---|---|
| Intercept (Standard) | 3.36 [2.73, 3.99] | 0.32 | 10.39 | 136.37 | <.0001 |
| Repeated group | 0.25 [-0.61, 1.11] | 0.44 | 0.57 | 117. | .56 |
| Gradual group | 0.36 [-0.50, 1.22] | 0.44 | 0.82 | 117. | .40 |
| Interview sessions | 0.16 [0.00, 0.32] | 0.08 | 1.87 | 239. | .06 |
| **Random effects** | ***SD*** | | | | |
| Participants | 1.37 | | | | |

Note: Coefficients are displayed with 95% CIs.

**Hypothesis 3.** With this hypothesis, we aimed to assess whether false memory implantation of repeated events would be facilitated by the method we employed in the Gradual group. Hence, this model compared the participants' self-rating for each occurrence of the false narrative (up to three) in the Repeated and Gradual groups for

all three interviews. Tables 7 and 8 show that, against our prediction, the participants in the Gradual group did not tend to provide higher recollection or belief ratings compared with the Repeated group. Moreover, the participants' ratings of both recollection and belief tended to increase significantly over the three interview sessions. Unsurprisingly, the second and third most vivid occurrences received significantly lower recollection and belief ratings compared with the first.

**Table 7.** Linear Mixed-Effect Model Results for False Recollection Across Repeated and Gradual Conditions.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***t*** | ***df*** | ***p* value** |
|---|---|---|---|---|---|
| Intercept (Repeated) | 1.72 [1.34, 2.10] | 0.19 | 8.92 | 97.01 | <.0001 |
| Gradual group | 0.08 [-0.42, 0.60] | 0.25 | 0.33 | 77.44 | .73 |
| Interview sessions | 0.22 [0.13, 0.32] | 0.04 | 4.57 | 550.71 | <.0001 |
| Second most vivid occurrence | -0.45 [-0.64, -0.28] | 0.09 | -4.94 | 550.71 | <.0001 |
| Third most vivid occurrence | -0.58 [-0.77, -0.41] | 0.09 | -6.36 | 550.71 | <.0001 |
| **Random effects** | ***SD*** | | | | |
| Participants | 1.09 | | | | |

**Table 8.** Linear Mixed-Effect Model Results for False Belief Across Repeated and Gradual Conditions.

| **Fixed effects** | **Unstandardised coefficient (*b*)** | ***SE*** | ***t*** | ***df*** | ***p* value** |
|---|---|---|---|---|---|
| Intercept (Repeated) | 3.60 [3.05, 4.15] | 0.28 | 12.817 | 98.38 | <.0001 |
| Gradual group | 0.09 [-0.65, 0.83] | 0.37 | 0.24 | 78.06 | .80 |
| Interview sessions | 0.19 [-0.05, 0.34] | 0.07 | 2.61 | 551.31 | <.01 |
| Second most vivid occurrence | -1.47 [-1.75, -1.21] | 0.13 | -10.83 | 551.31 | <.0001 |
| Third most vivid occurrence | -1.87 [-2.15, -1.61] | 0.13 | -13.75 | 551.31 | <.0001 |
| **Random effects** | ***SD*** | | | | |
| Participants | 1.59 | | | | |

## Discussion

Since false memory implantation studies have exclusively focused on single events, scholars have criticised such studies for potentially not applying to cases in which people allege multiple repeated instances of an offense (Blizard & Shaw, 2019; Brewin & Andrews, 2016). However, this study demonstrates that it is indeed possible to form false autobiographical memories of repeated events under laboratory conditions. Moreover, although we had expected the rate of false memory implantation to be lower in the Repeated and Gradual groups, surprisingly, the participants in the Repeated and Gradual groups formed false memories approximately as often and as vividly as the participants in the Standard group.

Many of our everyday experiences are repeated routine events. Without details that make an occurrence of a repeated event stand out, all occurrences tend to collapse

together into one gist or schema (Rubínová et al., 2020). For this reason, people may have an intuitive meta-memorial understanding that events they should have a schematic memory for are events they have experienced repeatedly. In the absence of such a schema, a person might doubt whether a suggested repeated event did in fact occur. Thus, we expected the participants to be more sceptical of the suggestion that they had experienced a false event repeatedly. This prediction was not borne out; instead, the participants in the Repeated and Gradual groups seemed no more sceptical than those in the Standard group.

Perhaps these results were due to people's intuitive understanding of how difficult it is to retrieve accurate details of repeated experiences. Because people may be intuitively aware of the poor quality of their memories of repeated events (e.g. the inability to differentiate between specific episodes and the inability to retrieve a single episode), their suspicion towards the veracity of the false narrative may have been mitigated. Another possibility has to do with the source of script memory. Scripts can be learned vicariously (Slackman & Nelson, 1984). That is, personal experience is not necessary to acquire a script for an event. Given that scripts are not tightly linked to specific episodes, it may be easy to lead a person to believe in a script they acquired through a source other than their direct experience.

As this is the first study investigating false memory implantation for repeated autobiographical events, we believe that our findings should be generalised with caution beyond the point of proof of concept that false memories of repeated events can be implanted. At this point, we cannot speak to the particular moderators that are likely to exist for false memory implantation for repeated events. Different types of false memory are not necessarily correlated (Chapter 3; Falzarano & Siedlecki, 2019; Ost et al., 2013; Zhu et al., 2013). Despite being elicited via the same method (i.e. the lost-in-the-mall-paradigm), it could be the case that false memories of repeated events and false memories of single events are elicited due to distinct mechanisms. Hence, understanding false memory moderators for single and repeated events may be a way to determine the extent to which studies on false memory implantation for single events might be generalised to cases on possible false memories of repeated events.

***Limitations and Future Research***

The false memory rate in our study for the Standard group was not as high as the false memory implantation rates reported in previous studies (Scoboria et al., 2017). Nevertheless, our results are not outside of the typical range (see Table A1). Differences between the implanted events in previous studies (e.g. positive versus negative events) and the way in which implanted false memories were scored (see Shaw & Porter, 2015; Wade et al., 2018) may play an important role in determining the rate of false memory formation. Importantly, to the best of our knowledge, the current experiment is one of the few experiments that used different methods to score false memories (i.e. self-reports

and researcher ratings; see Otgaar et al., 2013). Arguably, having two different ways to measure false memories reports, instead of only one, advocates for the reliability of our results.

Despite efforts to manufacture false events that share possible negative features with sexual abuse, false memory implantation studies share a limitation regarding the external validity of such false events. Unlike laboratory conditions, sexual abuse cases involve highly distressing events with personal relevance for the victims (for contextualisation on the complexity and effects of sexual abuse for victims, see: Marx & Sloan, 2002; Murrey et al., 2015; Schönbucher et al., 2012; Wolfe et al., 1994). There was no clear motivation for the participants in the current experiment to try to remember the episode we were attempting to implant – a limitation shared by previous studies on false memory implantation (see Table A1).

Sometimes, alleged victims who have had false memories of abuse eventually realise that their abusive memories were not authentic (Maran, 2010). When explaining their experience, they occasionally state that they were motivated to form such false memories because they wanted to resolve personal issues that could be explained by, for example, having experienced abuse (Lindsay & Read, 1994; Maran, 2010). However, 'motivation to recall' is a challenging feature for scientists to ethically include in false memory laboratory studies.

Another possible limitation concerns the fact that our research assistants, who interviewed the participants, were familiar with the design and hypotheses of our study. In such a logistically complex study, the interviewers must at least be familiar with the design and procedures. Even if we had attempted to fully blind them, it would have been obvious that all the participants were being told that they had lost an elephant cuddly toy during their childhood, making it easy to infer the purpose of that procedure. However, if the lack of blinding affected our results, it might be expected that our results would be more in line with our hypotheses. The fact that we found opposite results speaks against the severity of this threat to validity. Furthermore, the researchers who rated the memory reports were blind to the hypotheses. Overall, therefore, it is likely that our results were not severely affected by the interviewers' prior knowledge of the hypotheses of the current experiment.

Finally, our approach of letting the participants come up with the events of the general false experience we attempted to implant was necessary at this point in order for our method to be comparable with the classic implantation method. We believe, however, that future research on the implantation of false memories of repeated events should move from our methodology towards something more specific when implanting false memories of repeated events. For instance, a future study could define the exact episodes of the same experience to be implanted and then guide participants through them, which would be likely to increase the false memory rates for repeated events.

## Conclusion

This experiment offers the first demonstration that, under some circumstances, false memories of repeated autobiographical events can be implanted, in much the same way as false memories of single events can be implanted. Furthermore, if our findings can be generalised to repeated victimisation of sexual abuse, false memories of repeated events could have a powerful impact on people's lives – especially when such false memories answer relevant questions people may have about their own life (e.g. why they are experiencing depression or anxiety). When entering the courtroom, memory distortions may have high costs, often rendering years of emotional distress to the parties involved, and, ultimately, the imprisonment of an innocent person. Although this experiment cannot provide an estimate of the extent to which false memories of repeated events play a role in the legal system, it shows that they are possible. As it is clear that the potential consequences of false memories of repeated events are great, such false memories are a psychological phenomenon worthy of future attention in research.

# Appendix A

**Table A1.** Memory Implantation Method Studies Focused on Single Events in Adult Participants.

| Authors | False event | Type of stimulus | Guided imagery | Sample | Rates of any level of false memory formation | Additional information |
|---|---|---|---|---|---|---|
| Loftus and Pickrell (1995) | Being lost in a shopping mall or big department store | Participants read a description booklet | No | 24 adults | 20% | False narrative was adapted to be coherent with the participants' life |
| Hyman et al. (1995) Experiment 1 | Birthday with pizza and a clown visit/ overnight at the hospital | Researchers narrated the event | No | 20 adults | 20% | Selected 1 of the 2 events according to the participant's life; emotion valence between subjects |
| Hyman et al. (1995) Experiment 2 | Spilling punch at a wedding/ fire sprinkler being activated at a shop/ releasing the hand break of a car, causing a small crash | Researchers narrated the event | No | 51 adults | 26% | Selected 1 of the 3 events according to the participant's life; participants could add their own emotional interpretation of the false events |
| Hyman and Pentland (1996) | Spilling punch at a wedding | Researchers narrated the event | Yes (1 condition) | 65 adults | 12% without guided imagery 36% with guided imagery | Guided imagery between subjects |
| Pezdek et al. (1997) Experiment 1 | Doing something embarrassing at a religious ceremony (Catholic or Jewish) | Participants read a description booklet | No | 51 high school students (Catholic and Jewish) | 6% implausible 22% plausible | Plausibility between subjects |
| Pezdek et al. (1997) Experiment 2 | Being lost in a shopping mall/ receiving a rectal enema | Researchers narrated the event | No | 20 adolescents | 0% lower plausibility 15% high plausibility | Plausibility and script knowledge between subjects |
| Hyman and Billings (1998) | Spilling punch at a wedding | Researchers narrated the event | Yes | 66 adults | 27% | Use of individual differences measurements |
| Porter et al. (1999) | Undergoing a medical procedure/ getting lost/ being harmed by another child/ being attacked by an animal/ indoor accident/ outdoor accident | Researchers narrated the event | Yes | 77 adults | 56% | Selected 1 of the 6 events according to the participant's life; comparison of real, implanted and fabricated memories |
| Heaps and Nash (2001) | Life Events Inventory (LEI) | Researchers narrated the event | Yes | 63 adults | 37% | Selected 1 of the 42 LEI events according to the participant's life |

**Table A1.** Continued.

| Authors | False event | Type of stimulus | Guided imagery | Sample | Rates of any level of false memory formation | Additional information |
|---|---|---|---|---|---|---|
| Wade et al. (2002) | Hot air balloon ride | Participants viewed doctored photographs | Yes | 20 adults | 50% | Interviews followed a modified version of the step-wise procedure |
| Lindsay et al. (2004) | Putting Slime (a gelatinous toy) on the teacher's desk in Grades 1 or 2 | Researchers narrated the event, presenting or not presenting (real) photographs | Yes | 45 adults | 46% without photograph<br>78% with photograph | Presentation of a real school photograph between subjects |
| Garry and Wade (2005) | Hot air balloon ride | Participants read a description booklet containing or not containing (doctored) photographs | Yes | 44 adults | 50% with doctored photograph<br>82% with verbal narrative | Doctored photograph between subjects |
| Ost et al. (2005) | Going to the hospital/ getting lost/ family holiday/ birthday/ attending a wedding/ winning a contest/ death of a friend or relative/ accident with friend or family | Researchers narrated the event | No | 31 adults | 23% | Selected 1 of the 8 events according to the participant's life |
| French et al. (2006) | Hot air balloon ride | Participants accessed personalised events on a password-protected webpage through photographs or description | No | 58 adults | 24% | Investigated the influence of discussing false events with family members over false memory rates; for the false narrative, doctored photographs were used |
| Desjardins and Scoboria (2007) | Putting Slime (a gelatinous toy) on the teacher's desk in Grades 1 or 2, with or without the addition of self-relevant details and/or specific details | Participants read a description booklet | Yes | 44 adults | 68% with self-relevant details<br>36% without self-relevant details | Used the Autobiographical Belief and Memory Questionnaire (ABMQ); details: self- relevant and specific, self- relevant only, specific only, and no self- relevant or specific details were conditions between subjects |

**Table A1.** Continued.

| Authors | False event | Type of stimulus | Guided imagery | Sample | Rates of any level of false memory formation | Additional information |
|---|---|---|---|---|---|---|
| Qin et al. (2008) Experiment 1 | Birthday party at McDonalds/ getting lost/ receiving a rectal enema | Participants read a text description | Yes (1 condition) | 119 adults | Varies between participants according to condition: 36%–16% | Selected 1 of the 2 events according to the participant's life; used warning about false memory formation between subject |
| Strange et al. (2008) | Hot air balloon ride | Researchers narrated the event | No | 105 adults | 38% 2 years old 19% 10 years old | Age difference in which the false event occurred between subjects (2 years old vs. 10 years old) |
| Wade et al. (2010) | Hot air balloon ride | Participants read a description booklet and viewed a photograph booklet | Yes | 53 adults | 67% narrative first 41% doctored photograph first | Order of presenting doctored photograph (first or last) between subjects |
| Short and Bodner (2011) | Researchers selected plausible events for each participant that had happened to other participants in the study | Researchers narrated the event | Yes | 34 adults | 41% | Used the Judgment of Memory Characteristics Questionnaire (JMCQ) and the Cognitive Interview; used actual events, fabricated events and suggested events |
| Hessen-Kayfitz and Scoboria (2012) | Hot air balloon ride | Participants viewed a booklet of doctored photographs with a brief description | Yes | 82 adults | Varies between participants according to condition: 47%–24% | Used the Autobiographical Belief and Memory Questionnaire (ABMQ); presence of salient familiar details and absence of salient familiar details between subjects |
| Otgaar et al. (2013) Experiment 1 | Hot air balloon ride | Researchers narrated the event | Yes | 89 adults | 36% | Used the Memory Characteristic Questionnaire (MCQ) and Memory or Belief Form (MBF); also investigated non-believed memories |
| Shaw and Porter (2015) | Committing assault; committing assault with a weapon; committing theft; self-injury; being attacked by a dog; losing a lot of money | Researchers narrated the event | Yes | 60 adults | 70% criminal false memories 77% non-criminal false memories* | Rates were taken explicitly from self-reports. Criminal and non-criminal false memories between subjects |

*False memory has been operationalised differently across published studies. Furthermore, due to the unusually high false memory rates in this study, Wade and colleagues (2018) recoded the raw data and found rates to be between 26% and 30%.





Chapter V

**Table A2.** Scoring Definition for False Memory Categories.

| Category | Acceptance level | Elaboration level | Imagery level |
|---|---|---|---|
| Robust False Memory | High | Moderate or higher | Moderate or higher |
| Full False Memory | Moderate | Moderate or higher | Moderate or higher |
| Partial False Memory | Moderate | Low or higher | Low or higher |
| Accepted False Memory | Moderate | - | - |
| No False Memory | Low | - | - |
| Rejected False Memory | None | - | - |

# Appendix B

**Table B1.** Nationalities of Participants.

| Country | N | % |
|---|---|---|
| Germany | 28 | 23.3 |
| Netherlands | 22 | 18.3 |
| Belgium | 6 | 5 |
| Multiple nationalities | 5 | 4.2 |
| United States | 5 | 4.2 |
| Greece | 4 | 3.3 |
| Indonesia | 4 | 3.3 |
| Spain | 4 | 3.3 |
| Bulgaria | 3 | 2.5 |
| Finland | 3 | 2.5 |
| Ireland | 3 | 2.5 |
| Italy | 3 | 2.5 |
| Portugal | 3 | 2.5 |
| Brazil | 2 | 1.7 |
| Cyprus | 2 | 1.7 |
| France | 2 | 1.7 |
| Hungary | 2 | 1.7 |
| India | 2 | 1.7 |
| Mexico | 2 | 1.7 |
| Poland | 2 | 1.7 |
| United Kingdom | 2 | 1.7 |
| Canada | 1 | 0.8 |
| China | 1 | 0.8 |
| Ethiopia | 1 | 0.8 |
| Iran | 1 | 0.8 |
| Luxemburg | 1 | 0.8 |
| Norway | 1 | 0.8 |
| Peru | 1 | 0.8 |
| Sweden | 1 | 0.8 |
| Taiwan | 1 | 0.8 |
| Turkey | 1 | 0.8 |
| Zimbabwe | 1 | 0.8 |

**Table B2.** Participants' Gender.

| **Gender** | | |
|---|---|---|
| | **Frequency** | **%** |
| Female | 94 | 78.3 |
| Male | 26 | 21.7 |

**Table B3.** Participants' Age, Means and Median.

| **Age** | |
|---|---|
| N | 120 |
| M | 22.28 |
| Mdn | 21.00 |
| SD | 3.495 |
| Kurtosis | 0.534 |
| Minimum | 18 |
| Maximum | 32 |

**Table B4.** Frequency Distribution of Participants' Ages.

| **Age** | | |
|---|---|---|
| **N** | **Frequency** | **%** |
| 18 | 4 | 3.3 |
| 19 | 24 | 20.0 |
| 20 | 24 | 20.0 |
| 21 | 13 | 10.8 |
| 22 | 9 | 7.5 |
| 23 | 8 | 6.7 |
| 24 | 11 | 9.2 |
| 25 | 8 | 6.7 |
| 26 | 5 | 4.2 |
| 27 | 2 | 1.7 |
| 28 | 2 | 1.7 |
| 29 | 1 | 0.8 |
| 30 | 4 | 3.3 |
| 31 | 3 | 2.5 |
| 32 | 2 | 1.7 |

**Table B5.** Participants' Marital Status.

| Marital Status | | |
|---|---|---|
| | Frequency | % |
| Single | 117 | 97.5 |
| Married | 2 | 1.7 |
| Divorced | 1 | 0.8 |

**Table B6.** Participants' Educational Status.

| **Educational Status** | | |
|---|---|---|
| | **Frequency** | **%** |
| Missing | 1 | 0.8 |
| Not specified | 1 | 0.8 |
| Bachelor's degree | 76 | 63.3 |
| Master's degree | 39 | 32.5 |
| PhD | 3 | 2.5 |

# Appendix C

**Table C1.** Belief and Recollection after Debriefing.

| | **Rating** | **Number of observations** | **%** |
|---|---|---|---|
| Belief after debriefing | 1 | 72 | 60 |
| | 2 | 21 | 17.5 |
| | 3 | 4 | 3.3 |
| | 4 | 9 | 7.5 |
| | 5 | 4 | 3.3 |
| | 6 | 7 | 5.8 |
| | 7 | 3 | 2.5 |
| | 8 | - | - |
| Recollection after debriefing | 1 | 90 | 75 |
| | 2 | 13 | 10.8 |
| | 3 | 5 | 4.2 |
| | 4 | 4 | 3.3 |
| | 5 | 6 | 5 |
| | 6 | 2 | 1.7 |
| | 7 | - | - |
| | 8 | - | - |

Note: Table C1 contains the frequency and percentage of the participants' self-ratings for the false narrative after learning that they did not actually experience it. The rating was the same as that used for the stories during the interview sessions, in which 1 stood for 'this did not happen to me' or 'I do not recall this event' and 8 stood for 'I am completely confident that this happened to me' or 'I remember this event completely'.

**Table C2.** Counts of Memory Reports per Category of False Memory Formation across the Three Interview Sessions.

| **Standard Condition *n* = 39** Up to three measurements per participant (*n* = 117) | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Interview** | **Rejected False Memory** | **No False Memory** | **Accepted False Memory** | **Partial False Memory** | **Full False Memory** | **Robust False Memory** | **Total** |
| 1 | 4 (10.2) | 12 (30.7) | 18 (46.1) | 1 (2.5) | 2 (5.1) | 2 (5.1) | 39 |
| 2 | 5 (12.8) | 8 (20.5) | 16 (41.0) | 6 (15.3) | 1 (2.5) | 3 (7.6) | 39 |
| 3 | 6 (15.3) | 11 (28.2) | 14 (35.8) | 4 (10.2) | 2 (5.1) | 2 (5.1) | 39 |
| **Repeated condition *n* = 41** Up to nine measurements per participant (*n* = 369) | | | | | | | |
| **Interview** | **Rejected False Memory** | **No False Memory** | **Accepted False Memory** | **Partial False Memory** | **Full False Memory** | **Robust False Memory** | **Total** |
| 1 | 13 (10.5) | 81 (65.8) | 24 (19.5) | 3 (2.4) | 1 (0.8) | 1 (0.8) | 123 |
| 2 | 11 (8.9) | 82 (66.6) | 17 (13.8) | 6 (4.8) | 6 (4.8) | 1 (0.8) | 123 |
| 3 | 9 (7.3) | 72 (58.5) | 28 (22.7) | 8 (6.5) | 3 (2.4) | 3 (2.4) | 123 |
| **Gradual Condition *n* = 38** Up to seven measurements per participant (*n* = 266) | | | | | | | |
| **Interview** | **Rejected False Memory** | **No False Memory** | **Accepted False Memory** | **Partial False Memory** | **Full False Memory** | **Robust False Memory** | **Total** |
| 1 | 9 (25) | 0 | 20 (55.5) | 4 (11.1) | 2 (5.5) | 1 (2.7) | 36 |
| 2 | 15 (13.1) | 45 (39.47) | 34 (29.8) | 7 (6.1) | 12 (10.1) | 1 (0.8) | 114 |
| 3 | 11 (9.6) | 58 (50.8) | 30 (26.3) | 4 (3.5) | 10 (8.7) | 1 (0.8) | 114 |

Note: Values are frequency counts, with percentages in parentheses. The cases displayed in the table do not correspond to the number of participants, but rather to all measurements from each participant.

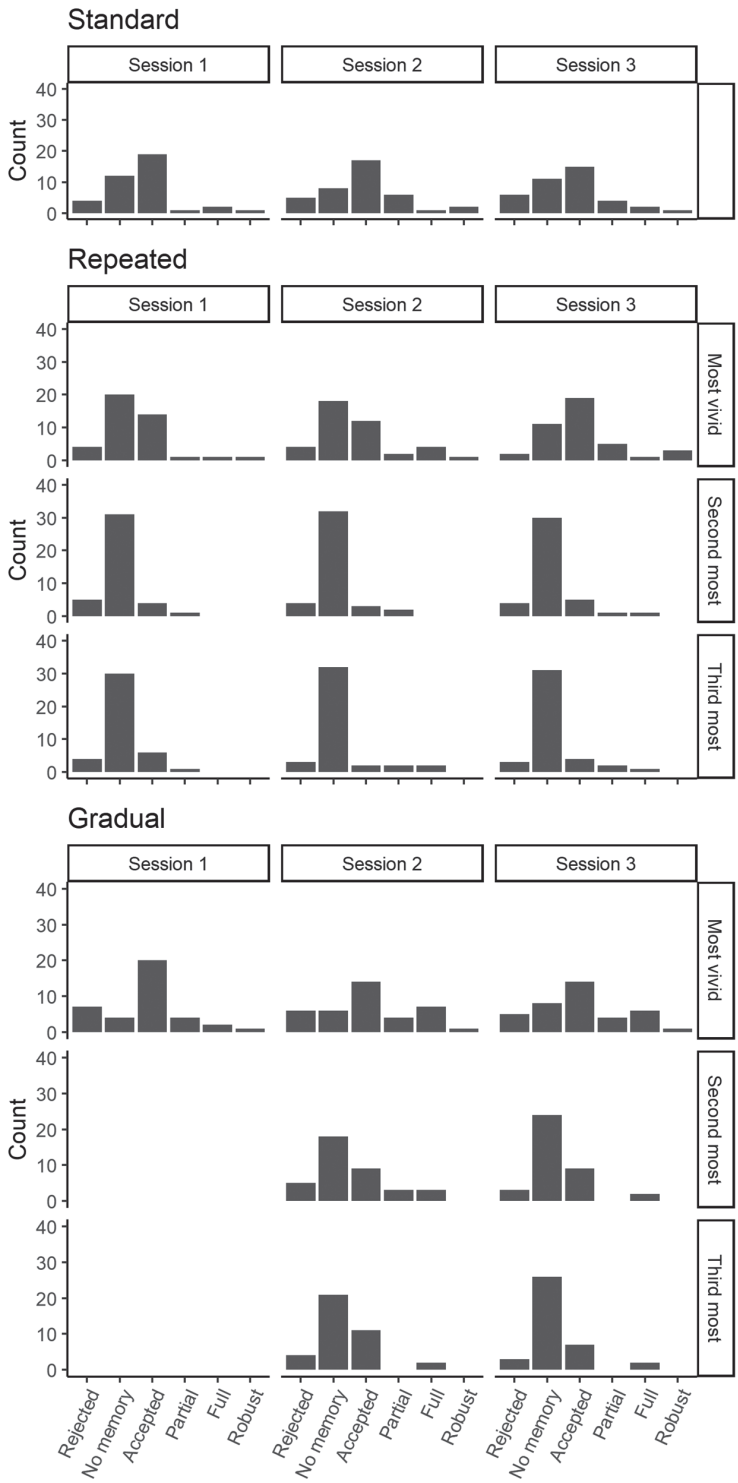
**Table C3.** Participants Describing Two or More Occurrences of the False Narrative Classified by the Researcher as Accepted, Partial, Full or Robust False Memory.

| **Repeated Condition *n* = 10** | | | | |
|---|---|---|---|---|
| **Participant** | **Occurrence** | **Interview Session 1** | **Interview Session 2** | **Interview Session 3** |
| 1 | 1 | - | - | Accepted |
| | 2 | - | - | Accepted |
| | 3 | - | - | Accepted |
| 2 | 1 | Robust | Accepted | Accepted |
| | 2 | - | Accepted | - |
| | 3 | - | - | - |
| 3 | 1 | Accepted | Full | Accepted |
| | 2 | Accepted | Partial | Full |
| | 3 | Accepted | Full | Accepted |
| 4 | 1 | Accepted | Accepted | Accepted |
| | 2 | - | - | Accepted |
| | 3 | - | - | - |
| 5 | 1 | Accepted | Accepted | Accepted |
| | 2 | Accepted | Accepted | Accepted |
| | 3 | Accepted | Accepted | Accepted |
| 6 | 1 | Accepted | Accepted | Accepted |
| | 2 | Accepted | Accepted | Accepted |
| | 3 | Accepted | Accepted | Accepted |
| 7 | 1 | Partial | Partial | Partial |
| | 2 | Partial | Partial | Partial |
| | 3 | Partial | Partial | Partial |
| 8 | 1 | Accepted | Accepted | Accepted |
| | 2 | Accepted | Partial | Partial |
| | 3 | - | - | - |

**Table C3.** Continued.

| **Repeated Condition *n* = 10** | | | | |
|---|---|---|---|---|
| **Participant** | **Occurrence** | **Interview Session 1** | **Interview Session 2** | **Interview Session 3** |
| 9 | 1 | Accepted | - | - |
| | 2 | Accepted | - | - |
| | 3 | Accepted | - | - |
| 10 | 1 | Accepted | Full | Accepted |
| | 2 | - | - | Full |
| | 3 | - | - | - |
| **Gradual Condition *n* = 14** | | | | |
| Participant | **Occurrence** | **Interview Session 1** | **Interview Session 2** | **Interview Session 3** |
| 1 | 1 | Accepted | Full | Partial |
| | 2 | - | Accepted | Accepted |
| | 3 | - | Accepted | Accepted |
| 2 | 1 | Accepted | Full | Accepted |
| | 2 | - | Accepted | Accepted |
| | 3 | - | - | Accepted |
| 3 | 1 | Accepted | Full | Full |
| | 2 | - | Full | Full |
| | 3 | - | Full | Full |
| 4 | 1 | Accepted | Accepted | Accepted |
| | 2 | - | Accepted | Accepted |
| | 3 | - | Accepted | Accepted |
| 5 | 1 | Accepted | Accepted | Accepted |
| | 2 | - | Accepted | Accepted |
| | 3 | - | Accepted | - |
| 6 | 1 | Accepted | Accepted | Partial |
| | 2 | - | Partial | - |
| | 3 | - | Accepted | - |
| 7 | 1 | Accepted | Accepted | Full |
| | 2 | - | Partial | Full |
| | 3 | - | Accepted | Full |
| 8 | 1 | Accepted | Full | Full |
| | 2 | - | Accepted | Accepted |
| | 3 | - | Accepted | Accepted |
| 9 | 1 | Accepted | Accepted | Accepted |
| | 2 | - | Accepted | Accepted |
| | 3 | - | Accepted | Accepted |
| 10 | 1 | Accepted | Accepted | - |
| | 2 | - | Accepted | - |
| | 3 | - | Accepted | - |
| 11 | 1 | Accepted | Full | Accepted |
| | 2 | - | Full | Accepted |
| | 3 | - | Accepted | Accepted |
| 12 | 1 | Accepted | Accepted | Accepted |
| | 2 | - | Accepted | Accepted |
| | 3 | - | Accepted | Accepted |
| 13 | 1 | Partial | Full | Full |
| | 2 | - | Partial | - |
| | 3 | - | Full | - |
| 14 | 1 | Accepted | Full | Accepted |
| | 2 | - | - | Full |
| | 3 | - | - | - |

**Figure C1.** False Memory Formation over Time for Each Group.



# Chapter VI

## General Discussion

In his autobiography, Nobel Laureate Gabriel García Márquez noted how memory shapes our lives: 'Life is not what one lived, but what one remembers and how one remembers it in order to recount it' (García Márquez, 2014, p. 6). His words contain both an idea of memory's centrality to human experience, and the notion that memory is reconstructive and is dependent on subjective perception rather than on recordings of a factual truth. Indeed, memory facilitates the creation and maintenance of individual and collective identity (Van Dyke & Alcock, 2003). Memories allow us to travel through our biographical experiences, attributing meaning to the present and future (Tulving, 1985).

Memories are not an exact copy of our past experiences. By retrieving memories, we are inevitably reconstructing them (Bartlett, 1932; Roediger et al., 2001). When memory errors – also known as false memories – are part of legal cases, they may concern entire autobiographical experiences that happen more than once. Furthermore, they are sometimes related to historical child sexual abuse cases (Classen et al., 2005). But as false memory reports of repeated abuse can be as vivid and rich in details as true memory reports of repeated abuse, it is difficult to single them out, which makes such cases particularly difficult to investigate (Loftus & Pickrell, 1995).

False autobiographical memories of repeated events can have far-reaching consequences (e.g. false accusations and wrongful convictions). Due to observed characteristics of cases of potential formation of false memories of repeated events (e.g. McMartin Preschool case), the credibility of sexual abuse statements as well as child memory performance in a co-witness condition are related and relevant topics within the theme of false memory of repeated events. Hence the main goal of my thesis was to introduce a novel case of potential formation of false memories of repeated events and empirically investigate the matters of memory conformity, credibility of reports of repeated sexual abuse, and finally formation of false autobiographical memories of repeated events.

## The Layout of This Thesis and the Summary of Chapters

The work comprising this thesis examined different aspects of the issue of false memories of repeated events within a child sexual abuse context. Instead of utilising a thesis format, in which one study serves as the basis for a chain of methodologically equivalent follow-up studies, I used different procedures to investigate related topics to false memory of repeated event cases. With this thesis, my ambition was ultimately to offer useful contributions to the debate of false memories of repeated autobiographical events cases and practical information for practitioners of the legal system with regard to such cases.

**Chapter I** began with an introduction of the fundamental research and theories related to individual chapters of the current thesis. Next, I grounded the central theme of this thesis using two legal cases that exemplify the possible consequences of false memories

of repeated events in the legal field. Meredith Maran's case is a classic example of false memory formation, in which – after several therapeutic sessions – a patient comes to remember episodes of child sexual abuse that she had never seriously considered before (Coons, 2008). Maran's case is one of many cases that took place in the 1980s in the United States, in which women came to believe that they had been sexually abused several times by close relatives, but later realised those to be false memories (Maran, 2010). Naturally, Maran's and other similar cases had a direct and significant impact not only on the alleged victims, but also on other people's lives. Many families fell apart, never to reconnect (see Yuhas, 2021). Many parents who had been accused of sexually abusing their children at that time created associations and groups to support one another through the prosecutions of the respective cases (Maran, 2010).

The McMartin Preschool case concerns potential false reports of child sexual abuse revictimisation and the potential formation of false memories as a product of suggestive techniques being used in investigative interviews with children. The case lasted for almost a decade, temporarily incarcerating the accused – who consequently lost their jobs – and costing the state of California, USA, millions of dollars (Young, 2008). The McMartin Preschool case had such severe consequences that, despite the lack of certainty regarding whether the McMartin children had formed false memories, it is a robust illustration of the impact that memory fallibility can have in the legal system.

In Chapter II**,** I presented a review with clear parallels to the Maran and McMartin Preschool cases: the Jakarta International School (JIS) case from 2014. In the JIS case, three children reported graphic descriptions of repeated sexually abusive acts, allegedly performed by their school staff. At first, the children vehemently denied that any abuse had ever taken place. Strikingly, after about a month of investigation in which the children were subjected to an array of external suggestive sources, the children confirmed that they had been sexually abused by the JIS staff. Moreover, the JIS children's statements changed gradually in terms of the complexity and frequency with which they had allegedly been abused. Chapter II refutes the common belief that legal cases in which false memories are fostered of repeated child sexual abuse are a relic of the past, thus confirming the continuing need for the research presented herein. Furthermore, the JIS case provides a relatively recent illustration of how certain interviewing techniques might trigger statements tainted with false information (e.g. Poole & White, 1991).

Chronologically speaking, Chapter II comprises the last project I developed for my thesis, however, the subsequent chapters can be seen as drawing from salient issues in the JIS case. In Chapter III, I investigated how the children in the JIS case could have contaminated one another's memories and, hence, their memory statements, by interacting with each another and sharing their impressions of the occurrences of sexual abuse they had allegedly suffered. In Chapter IV, I investigated whether children reporting revictimisation (as they did in the JIS and McMartin cases) would appear to be more credible to relevant parties than children reporting a unique account of child

sexual abuse. In Chapter V, starting from the assumption that, in the dubious day care abuse cases, the children not only may have falsely reported being sexually abused, but also formed actual false memories of the scenes they reported to remember, I investigated whether it would be possible to form false memories of entire autobiographical events that allegedly happened repeated times.

## Overview of the Main Findings of the Chapters in This Thesis

As mentioned earlier, Chapter II presents a case report analysing the most salient aspects of the JIS case, within the theme of child suggestibility regarding fostering false memories. By drawing on the legal psychological literature to assess the JIS case, I concluded that, despite the growing number of studies and recommendations to practitioners that have been published over the years, inadequate investigation techniques are still in use (Korkman et al., 2008; Sumampouw et al., 2019). Besides missing out on the optimal format to interview the children and hence preserve the integrity of their memories, the JIS case investigation team did not consider any alternative scenarios apart from the scenario that the children provided a truthful account of abuse (see **Chapter II**). The lack of evidence – such as other witnesses or DNA samples – to corroborate the children's statements makes the JIS case the stage of a cascade bias effect (Dror et al., 2017), which culminated with the incarceration of potentially innocent people.

The JIS case can also be related to archaic Swedish cases that happened in the 17$^{th}$ century, particularly in terms of one child attesting the involvement of other children, who were subsequently coerced to disclose false events (see Sjöberg, 1997). There is a clear indication that better practice would have prevented such detrimental outcomes in these cases. Regardless of the culpability of the accused in the JIS case, the use of appropriate techniques in the investigation, particularly for interviewing, would have increased the chances of collecting reliable statements from the children (Schreiber et al., 2006). The JIS case also highlights a flaw that must be improved in the dissemination of scientific findings to practitioners. Despite the array of publications addressing legal professionals' practice, there is never a guarantee that practitioners will access the research that is pertinent to their work.

It was necessary to demonstrate that false memory cases are still likely to occur, despite several decades of informative scientific studies aiming to prevent them. This brought me to the conceptualisation of the JIS case report in Chapter II. My argument is that case reports such as that in Chapter II are relevant to the praxis of legal professionals, as well as inspiring different directions for future research. The JIS case report can also be thought of as an introduction to the denser projects of the current thesis, as it anchors the somewhat abstract topics of Chapters III, IV, and V by providing a concrete illustration of the concepts they investigate.

Chapter III contains an empirical study at Dutch primary schools that investigates children's memory performance in a co-witness situation (memory conformity paradigm; Gabbert et al., 2002) and asks whether such a performance would correlate with the children's propensity to spontaneously form false memories (the DRM paradigm; Deese, 1959; Roediger & McDermont, 1995). That is, how prone are children to produce an unreliable statement if they do not receive external suggestions? The investigation showed that, compared with an older age group, children performed poorly in terms of producing a reliable eyewitness statement in a co-witness context. This finding means that, when children's statements are contaminated with suggestions from a peer who has witnessed the same event, children tend to conform and introject their peer's input, rather than holding on to their own memory.

Furthermore, corroborating other research on developmental reversal trends, we found that children performed better than an older group at reporting incorrect words from DRM paradigm lists (see Holliday et al., 2011). This result implies that, when children are not given suggestive information and are instead asked to freely recall and report what they can remember about what they have witnessed, children are less likely to provide inaccurate statements. Applying the results from Chapter III to the JIS case in Chapter II, I could argue that it is likely that the JIS children did not provide reliable statements to the police.

Regardless of the culpability of the accused in the JIS case, the JIS children did not have the opportunity to report what they freely recalled about the alleged occasions of sexual abuse. The questions directed to the three JIS children since the beginning of the investigation were already embedded with information that established a unique scenario in which the three boys – and, potentially, more children enrolled in the JIS – had been abused. In particular, having play dates and therapy sessions together allowed the three children to share information with one another. In the transcripts of the children's interviews, it is very clear that the children's version of what supposedly happened overlapped, as the children made direct references to versions of the sexual abuse stories that had been provided by their peers as they shared their stories among themselves.

Another striking aspect of the JIS and McMartin Preschool cases was that the children's perceived credibility was not affected by their claims of being revictimised by the JIS staff. The literature of credibility assessment shows that evaluators attest less credibility to a speaker who narrates a repeated event than to a speaker who narrates a single event (Connolly et al., 2008[a,b]). However, such studies did not use statements from alleged sexual abuse victims. Hence, via an online experiment in Chapter IV, we examined legal system professionals' and the general public's perceptions of the credibility of sexual abuse victims when narrating revictimisation versus single-occurrence abuse.

We also checked whether evaluators' perception of credibility would be affected by the victim's age and the time of disclosure. In terms of the credibility attributed

to victims reporting repeated sexual abuse versus victims reporting a single event of abuse, our results replicated previous findings in the credibility assessment literature (see Price & Connolly, 2004). That is, our participants found victims who reported repeated sexual abuse to be less credible than victims who disclosed one-time sexual abuse. In terms of the age of the victims, both adult and child victims were found to be equally credible, and there was no difference in terms of whether the adult victims reported current abuse or historical abuse.

The fact that our results are akin to previous research not based on sexual abuse vignettes does not necessarily indicate that the sexual abuse allegation theme of the statements is not relevant to other research in the more general field of false memories. For instance, we observed that biases, such as victim blaming, were present in the assessment of some practitioners when attesting the credibility of the victims' reports, which can be relevant for future research. Moreover, our findings do not support the events that took place in the JIS case and in other dubious day care abuse cases, such as the McMartin Preschool case. That is, the children in those cases were perceived as credible despite their allegations of repeated sexual abuse. In the JIS case, other elements might have biased the investigation to attribute high credibility to the children's accounts of sexual abuse revictimisation.

At first, the JIS children only reported one occasion on which they were allegedly sexually abused. As they received more suggestive information throughout the investigation, they also reported other occasions of sexual abuse. This gradual increase in the number of times the children stated that they had been sexually abused may have played a role in the credibility that was attested to the children. Another factor may have been that – starting at the onset of the first suspicions of sexual abuse for the first child involved in the case – the JIS case was broadcasted in the Indonesian media with a potential bias towards the veracity of the sexual abuse suspicions (see Hans & Dee, 1991).

Extrapolating beyond false allegations of sexual abuse, it is possible to speculate a scenario in which the JIS children may have formed false memories of repeated occurrences of child sexual abuse. Considering such a scenario, we wanted to investigate whether false memories of repeated events could be fostered under laboratory conditions. Such an investigation would also address a cluster of relevant critiques of the false memory implantation literature (e.g. Brewin & Andrews, 2017).

The literature has shown that false memories of an entire autobiographical event can be implanted into a person's memory (see Scoboria et al., 2017). However, to the best of our knowledge, until now memory scientists have not investigated the implantation of false memories of repeated events. Some scholars consider that an exclusive focus on the implantation of a single-frequency false autobiographical event posits an ecological fragility within the false memory implantation literature by making such research irrelevant to real-life situations. That is in fact the case, since sexual abuse is often a reoccurring offense for many victims (see Blizard & Shaw, 2019). In the laboratory

experiment in **Chapter V**, we adapted the lost-in-the-mall paradigm (Loftus & Pickrell, 1995) – also known as the implantation method – to assess whether false memories of repeated events could be formed.

Our experiment showed that false memories of repeated events can indeed be fostered under lab conditions. Moreover, it is not more difficult to implant false memories of repeated events than to implant false memories of single events. This demonstration that false memories of repeated events can be elicited bridges an important gap in the false memory implantation literature, while contributing to the debate of repressed memories as well (see Otgaar et al., 2019). Many scholars who criticise the ecological validity of false memory implantation work are proponents of the argument that a person can repress memories of traumatic experiences (see Blizard & Shaw, 2019). These scholars believe that people can completely forget being sexually abused and are later able to recover such memories with the aid of specific techniques.

The 'Memory Wars' is a current, yet old debate grounded in a popular discussion from the 1990s between memory scientists and clinical psychology scholars (see Otgaar et al., 2019). The specific line of argument of these clinical psychology scholars is that the research conducted on false memory implantation cannot be related to cases such as Maran's. Their justification lies in the fact that memory researchers focus on investigating the (false) formation of autobiographical events that happened only once. In Maran's case and similar cases, the alleged victims report revictimisation; hence, the scenario is not the same as that investigated in false memory implantation studies.

Ultimately, such scholars have argued that, because memory scientists had not demonstrated false memory formation for repeated autobiographical events, Maran's and similar cases were likely to be a product of recovered memories of true experiences (Blizard & Shaw 2019; Brewin & Andrews, 2017). Consequently, this argument could also be extended to the JIS case, which exhibits many of the features that are known to produce false reports and false memories. However, as mentioned in the discussion of Chapter V, our results contradict our predictions, which were founded on the theoretical framework presented in Chapter I. That is, considering theories such as schema theory (Friedman, 1979), fuzzy trace theory (Brainerd et al., 2008), and associative activation theory (Howe et al., 2009), we predicted that false memories of repeated events would be significantly more difficult to engender than false memories of single events. The fact that our results did not corroborate our prediction may indicate that other elements – such the form of social pressure within the dynamics of the interaction between an interviewer and interviewees – are interfering with the interviewees' confidence that a false suggested event did not occur to them, despite being unable to attribute a source for the event (see Ost et al., 2001).

## Practical and Theoretical Implications

Despite the variation in focus of investigation, the projects that coalesce in this thesis are relevant for legal practitioners, psychotherapists, and (of course) researchers. In Chapters I and II, we observed areas requiring improvement in psychotherapists' (the JIS case and the Maran case) and practitioners of the legal system' (the JIS case and the McMartin case) tools and protocols. Indeed, the current thesis contains four studies that can be informative for such professionals. The JIS case report illustrates the impact of suggestion from professionals who may be unaware of child suggestibility and memory fallibility. In the JIS case, the therapist was responsible for reinforcing the idea in the children that they had been sexually abused. Furthermore, the police officers that interviewed the children and conducted the investigation of the case contributed to the children's potentially false allegations. In the interviews with the children, the police officers consistently reinforced the children's parents' belief and never allowed the children to freely offer a possible version of what could have happened.

The Indonesian court has ruled that all the accused in the JIS case are guilty. However, this analysis of the JIS case, which considered the corpus of research investigating child suggestibility and false memories, revealed that the case is extremely relevant as an example of another dubious day care abuse case, akin to the McMartin Preschool case. Interestingly, the JIS case arose in a non-Westernised society, about 30 years after the sexual abuse hysteria of the 1980s. The existence of the JIS case certainly raises the question of how often false memory cases occur. Shaw and Vredevelt (2019) reported that the archive of the British False Memory Society in the United Kingdom has over 2500 cases of alleged false memories registered from 1994 to 2018. Furthermore, recent work indicates that it is not only in Indonesia that the legal system struggles with incorporating scientifically backed practices (Otgaar et al., 2019; Sumampouw et al., 2019).

Improvement in the practices used to interview children in such contexts might come from an understanding of children's optimal performance conditions when providing a statement. A corpus of research provides evidence that open questions allow children to freely recall what they have witnessed, resulting in reliable statements (Berg et al., 2019). Chapter III corroborates this line of research. Furthermore, practitioners can find information in Chapter III attesting that children are highly influenceable by social conditions – not only by adults, but also by their own peers. Hence, protecting children from such influences by using scientifically substantiated protocols might ensure that children can provide reliable statements (see Lamb et al., 2007). Children can be highly reliable when they are prevented from external contamination (Ceci & Bruck, 1993). However, when it is impossible to shield children from receiving input from others who might compromise the children's statements, it is relevant to acknowledge that the children's statements might be tainted with untrustworthy information (Principe et al., 2012, 2013).

Taken together, Chapters IV and V show relevant results for the practice of legal professionals. Chapter IV is educative about a possible bias legal professionals may exhibit in disregarding sexual abuse statements as having little credibility because the victims narrate multiple occasions of abuse. Sexual abuse is likely to happen to a victim repeatedly instead of once (Classen et al., 2005). Hence, such a bias towards repeated allegations of child sexual abuse can cripple investigations of true cases. Chapter V exemplifies an extreme scenario involving the (possible) formation false memories of repeated autobiographical events resulting from the use of suggestion and misinformation over time. Even when the participants were sceptical of the suggestions that they had experienced our study's false narrative, some came to at least believe in the interviewer's suggestions, if not produce their own false memories of the event(s), after being misinformed three times over a period of 3 weeks. The social influence exercised by our interviewers over the participants may be comparable to the social influence dynamic in a legal interview setting.

From a theoretical perspective, the current results might offer some new insights on how false memories for repeated events are formed. In previous research, when children were subjected to experiencing either a single or repeated event, children presented higher propensity to form false memories for repeated events rather than single events (Connolly & Lindsay 2001). In the light of fuzzy trace theory (Brainerd & Reyna 2002) or schema theories (Ghosh & Gilboa, 2014) such findings show that retrieving specific details of individual occurrences of the repeated event is difficult. When being unable to recall specific details about a given event, people rely on the overall idea or expectation they might have of the given event that would correspond to potentially all occurrences of the repeated event. Hence, a given repeated event can amplify susceptibility to forming false memories about the details of the event occurrences. This general finding, combined with the finding in Chapter 3 that children are less likely to form spontaneous false memories than adolescents, supports the tenets of fuzzy trace theory and schema/script theory.

However, for implanted false memories for repeated events, fuzzy trace theory could not provide a clear explanation for the results reported in Chapter V. Specifically, we found that false memories of single and repeated events were equally likely to be elicited. How can the latter finding be explained in light of the current theories? One option is that the used suggestion irrespective of referring to a single occurrence or repeated occurrence evoked a feeling of familiarity in participants in all experimental groups. That is, since the false narrative used was general with no specific details for each participant, participants might have experienced the single and repeated suggestion as equally familiar which made the false narrative equally likely to be implanted into memory. One way to examine a possible influence of familiarity into false memory implantation for repeated events is by replicating the work in Chapter V using a more specific false event in addition to asking participants for familiarity ratings of the false event.

Another way to assess results in the current thesis in light of theoretical implications would be to look into some of the effect sizes obtained in the current thesis. For instance, in Chapter 4, Table 2, we see an unstandardised coefficient of $b$ = -0.31 for the credibility attested to sexual abuse victims reporting revictimisation compared to sexual abuse victims reporting a single-occurrence abuse. Such effect of $b$ = -0.31 is relatively small given the six-point scale of this measure, indicating that there was not a large difference on how professionals of the legal system evaluated the two different statements. Furthermore, in Chapter V, when comparing our implantation rates back to the literature benchmarks for false memory implantation (single events) our rates of 25% (n = 39) were significantly lower than the overall implantation rate of 46% (n = 423) across previous studies (Scoboria et al., 2017). Despite these seemingly modest effects, phenomena in the legal system do not need to occur in large volumes or magnitudes to cause noteworthy distress. For example, the demonstration that false memories for repeated events can be implanted should be seen as a warning sign to professionals dealing with cases on revictimization where suggestion took place. Throughout my thesis, there are practical examples of such cases and how they have jeopardised family bonds and entire communities as well as costed substantial financial investment in lengthily investigations and trials (e.g. McMartin Preschool and Meredith Maran cases).

***Limitations and Future Research***

As a culmination of the work I performed during my doctorate, this thesis is an effort to contribute to the existing literature that examines practical issues psychology can address within the legal field. Chapter II, a legal case, is an appropriate illustration of the relevance of the particular themes discussed in this thesis. However, as with most of the knowledge researchers produce, there is a limitation on how much the current work will specifically and directly aid problems in the field. In past few years, I had the opportunity to occasionally meet legal professionals at different conferences I attended, including attorneys, forensic psychologists, and police officers.

When conversing with these practitioners, they explicitly spoke of their disappointment in the lack of dissemination of the knowledge we academics produce into the practical world. Indeed, if cases like the JIS continue to occur, it seems clear that insufficient practitioners are acquainted with the issues researchers have been studying in academia. Consequently, a critical issue for this thesis is whether its content will sufficiently reach those to whom it is most relevant. Education might be an effective strategy (see Sauerland & Otgaar, 2021). However, effectively prompting practitioners to undertake further education to improve their practice can be a strenuous task. Researchers and practitioners do not necessarily get along, as is hinted at by the wording used to name the debate on repressed and false memories (i.e. the "Memory Wars"). Hence, practitioners' self-will to educate themselves on scientifically grounded practices may not be enough, particularly given that each country has its own dynamics regarding

the extent to which public policy relies upon science.

Countries around the world have different regulations for the professions related to the central theme of this thesis (i.e. psychotherapists, social workers, and police officers). For instance, psychotherapists in Brazil have their practice regulated by a specific federal council (the Conselho Federal de Psicologia). When denounced by a third party for malpractice, psychotherapists in Brazil must respond to the accusations at the federal level; if found guilty, they may lose their license to practice clinical psychology, with a possibility of being prosecuted as well, depending on the severity of their actions (Processos Éticos, n.d.). Although it is outdated in terms of its taking research on false memory formation under consideration, the Conselho Federal de Psicologia in Brazil is an example of how policymaking could be a determinant in assuring that scientific findings are applied into professional praxis.

Some academics, including myself, aim to tighten their communication with practitioners by collaborating in projects that arise directly from practitioners' necessities and attending practitioners' conferences in order to share their contributions. However, restrictions in time and resources can prevent academics from extrapolating their research findings further than scientific outlets. For instance, in my 3-year PhD programme, there was not much I could do to disseminate my work beyond 'recommendation for practitioners' sections in my publications and a once-a-year presentation at a conference. I believe that academics' communication with practitioners should go beyond publications in scientific journals. Although a number of practitioners keep up to date with relevant scientific findings, the variation of socioeconomical structures around the world indicates that many practitioners lack the resources to stay up to date with scientific research. Different vehicles may be needed to establish an organic relationship between academics and those from whom we receive inspiration for our research questions and to whom we dedicate our work. This change in our culture remains to be achieved.

Moreover, the use of an action research modality, which is possible in the social sciences, can facilitate communication with practitioners (see Rowell et al., 2015). Action research ensures that there is a practical return to a community or society in exchange for the information participants lend researchers. Such a return often comes in the form of projects that improve the context in which the research took place. Furthermore, in the application of qualitative research tools within quantitative research, I see a channel to ease our communication with practitioners and the public. I advocate the use of qualitative research capacities not only to enrich our understanding of our own findings, but also to facilitate the clarification of discussions and themes that are often perceived as abstract by practitioners.

Thematic analyses are a common qualitative tool within the social sciences (Joffe, 2011). In **Chapter IV**, when comparing the perceived credibility of sexual abuse statements on a single sexual abuse occurrence versus multiple occurrences, we added an open-ended question asking our participants to justify their choice of one

statement over the other. Of course, the participants' answers allowed us to understand whether our manipulation worked; however, it also enriched our comprehension of the credibility assessment our participants had provided. To be specific, the participants' answers contained more themes than just 'single versus repeated'. For instance, the stereotypes the participants attributed to the victims of sexual abuse enlightened us on the fact that even trained legal practitioners may contribute to naturalising and disseminating oversimplified attributions of victims' behaviour, despite the amount of research addressing this issue.

Moreover, some of the practitioners of the legal system we tested seemed to connect the details of the sexual abuse statements we used with specific situations within their own cultural context. Such a strategy could hint at a possible lack of flexibility in acknowledging different mores around the world. As an example, one of our participants discredited the statement in which a child helped with home chores such as picking up items in the basement because, according to the participant in question, that would not happen within her or his own culture. Of course, overall, such a lack of flexibility is not a direct issue, since it is likely that legal practitioners will operate within their own cultural background. Nevertheless, it is possible for practitioners to be called to join international collaborations in which they might be confronted with diverse realities. Cultural differences are important aspects to be introjected within research with practitioners; thus, cross-cultural studies hold the potential to enhance our understanding of the differences that exist between countries in the assessment of sexual abuse credibility (see Anakwah et a., 2020; Sumampouw et al., 2021).

In the experiment in Chapter V, which focused on the implantation of false memories of repeated autobiographical events, we also asked the participants at the end of the study, to tell us whether they could identify which story was not part of their childhood, among the narratives we told them. As the interviews had been audio recorded, we coded in the participants' answers beyond the 'yes/no' option on the rating sheet. The inclusion of the participants' commentary on the false narrative allowed us to see that, even though the participants were able to recognise the false narrative most of the time, some still displayed literal verbal surprise when told the correct answer. Although this was not the focus of the work developed for Chapter V, the fact that some participants were surprised that the false narrative was not part of their childhood might be relevant for future research on related topics, such as assessing the deterioration of false memories of repeated events over time and non-believed memories (see Clark et al., 2012).

Future research might also benefit from drawing more direct inspiration from the situations and contexts in which the research needs to be applied. For instance, studies assessing the credibility of witness statements that narrate single versus repeated events could consider having the use of ethical adaptations of sexual abuse statements to be assessed by the participants, since this is one of the most relevant contexts of a credibility assessment. Research shows that people have various biases regarding sexual abuse

victims (see Burt, 1980; Check & Malamuth, 1983). Future research could expand on this topic by investigating how such stereotypes interfere with the credibility that is attributed to statements on repeated-frequency versus single-frequency sexual abuse.

In the field of false memory implantation, it has taken almost 30 years for an experiment to be conducted on the implantation of false memories of repeated events. As previously discussed, the lack of research addressing the formation of repeated false autobiographical events may have reinforced scholars' allegations discrediting the literature on false memory formation for autobiographical events. In view of a possible continuous critique of false memory implantation studies, future research could consider adjusting the implantation paradigm to investigate the formation of false memories of repeated events instead of single events. The study in Chapter V is the first to assess false memories of repeated events and, despite our diligence with the work conducted, we advocate that more studies are needed to substantiate this topic. For instance, Oeberst and colleagues (2021) have shown that implantation of false memories of a single autobiographical event can be reversed; yet it is not clear whether the same pattern would be observed for repeated false autobiographical memories. A practical example of reverting a false memory is Meredith Maran's case presented in Chapter 1.

Unlike scientific studies such as the one in Chapter V, in which participants underwent suggestive techniques only three times in total, cases such as Maran's are a product of years of therapy. Furthermore, it is difficult to ethically induce under laboratory conditions the motivations people may have to remember a spurious event. The women looking for therapeutic aid in the United States during the 1980s often had a strong motivation to understand why they had a bad relationship with their father or were depressed. Perhaps future research can also investigate the impact of motivation in forming false autobiographical memories.

## Conclusions

The literature is filled with studies probing the trustworthiness of memory. When memory fails and the authorship of a non-committed crime is mistakenly attributed to an innocent person in a legal statement, memory errors can have severe consequences. In the past three decades, scientists have examined the different characteristics of contemporary dubious child sexual abuse cases. Such studies are important to prevent the repetition of such detrimental cases, particularly if practitioners can become acquainted with these studies. This thesis presented six chapters with practical relevance that coalesce to address the issue of false memories of repeated occurrences of child sexual abuse. Furthermore, I advocate the adherence of initiatives to improve the communication between academics and practitioners, such as workshops, conferences and open channels of discussion even via social media outlets. Notwithstanding the effort academics already put into disseminating our work to those it is most relevant

for, I believe it is important to continue thinking of ways to facilitate communication between academia and praxis to prevent possible serious miscarriages of justice due to false memories while ensuring that cases of true memory receive the credibility they deserve.

# Appendices

References
Impact
Summary
Sammanfattning På Svenska
Acknowledgements
Curriculum Vitae

## References

*70% das vítimas são crianças e adolescentes: oito dados sobre estupro no Brasi*l. (2016, May 27). BBC New Brasil. https://www.bbc.com/portuguese/brasil-36401054

Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology*, *34*, 1358–1372. doi:10.1037/0012-1649.34.6.1358

Albright, T. D. (2017). Why eyewitnesses fail. *Proceedings of the National Academy of Sciences*, *114*, 7758–7764. doi: 10.1073/pnas.1706891114

Allan, K., & Gabbert, F. (2008). I still think it was a banana: Memorable "lies" and forgettable "truths." *Acta Psychologica, 127*, 299–308. doi: 10.1016/j.actpsy.2007.06.0

Anakwah, N., Horselenberg, R., Hope, L., Amankwah-Poku, M., & Van Koppen, P. J. (2020). Cross-cultural differences in eyewitness memory reports. *Applied Cognitive Psychology*, *34*, 504-515. doi: 10.1002/acp.3637

Anastasi, J. S., & Rhodes, M. G. (2008). Examining differences in the levels of false memories in children using child-normed lists. *Developmental Psychology*, *44*, 889–894. doi: 10.1037/0012-1649.44.3.889

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261–295. doi: 10.1016/s0022-5371(83)90201-3

Bardin, L. (1977). *Análise de conteúdo*. Edições 70.

Bardin, L. (2011). *Análise de conteúdo*. Edições 70.

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

Bass, E., & Davis, L. (1988). *The courage to heal: A guide for women survivors of child sexual abuse.* Perennial Library/Harper & Row Publishers.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi: 10.18637/jss.v067.i01.

Berg, R., Munthe-Kaas, H. M., Baiju, N., Muller, A. E., & Brurberg, K. G. (2019). The accuracy of using open-ended questions in structured conversations with children: A systematic review. *Norwegian Institute of Public Health. Retrieved from: https://www.fhi.no/en/publ/2019/The-accuracy-of-using-open-ended-questions-in-structured-conversations-with-children/*

Bernard-Bonnin, A. C. (2000). Diurnal enuresis in childhood. *Canadian Family Physician Medecin de Famille Canadien*, *46*, 1109–1115.

Bernstein, D. M., Pernat, N. L. M., & Loftus, E. F. (2011). The false memory diet: False memories alter food preferences. In: V. R. Preedy, R. R. Watson, & C. R. Martin (Eds), *Handbook of Behaviour, Food and Nutrition* (pp. 1645–1663). Springer.

Bernstein, D. M., Scoboria, A., Desjarlais, L., & Soucie, K. (2018). 'False memory' is a

linguistic convenience. *Psychology of Consciousness: Theory, Research, and Practice*, *5*, 161–179. doi: 10.1037/cns0000148

Blizard, R. A., & Shaw, M. (2019). Lost-in-the-mall: False memory or false defence? *Journal of Child Custody*, *16*, 20–41. doi: 10.1080/15379418.2019.1590285

Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and false memory: The context-content paradox. *Psychological Bulletin*, *142*, 1315–1351. doi: 10.1037/bul0000077

Brainerd, C. J. (1990). Issues and questions in the development of forgetting. *Monographs of the Society for Research in Child Development, 55*, 100–109. doi: 10.1111/j.1540-5834.1990.tb00287.x

Brainerd, C. J. (2013). Developmental reversals in false memory: A new look at the reliability of children's evidence. *Psychological Science*, 22, 335–341. doi: 10.1177/0963721413484468

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164-169. doi: 10.1111/1467-8721.00192

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

Brainerd, C. J., & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review*, *32*, 224–267. doi: 10.1016/j.dr.2012.06.008

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, *3*, 343–382. doi: 10.1037/0033-2909.134.3.343

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development*, *73*, 1363–1377. doi: 10.1111/1467-8624.0047710.1093/acprof:oso/9780195154054.001.0001

Brainerd, C. J., Holliday, R. E., Reyna, V. F., Yang, Y., & Toglia, M. P. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology*, *107*, 137–154. doi: 10.1177/0963721413484468

Brainerd, C. J., Stein, L. M., Silvera, R. A., Rohenkohl, G., & Reyna, V. F. (2016). How does negative emotion cause false memories? *Psychological Science*, *19*, 919–925. doi: 10.1111/j.1467-9280.2008.02177.x

Brewin, C. R., & Andrews, B. (2017). Creating memories for false autobiographical events in childhood: A systematic review. *Applied Cognitive Psychology*, *31*, 2–23. doi: 10.1002/acp.3220.

Browne, A., & Finkelhor, D. (1986). Impact of child sexual abuse: A review of the research. *Psychological Bulletin*, *99*, 66–77. doi.o: 10.1037/0033-2909.99.1.66

Bruck, M., & Ceci, S. (2004). Forensic developmental psychology. *Current Directions in Psychological Science*, *13*, 229–232. doi: 10.1111/j.0963-7214.2004.00314.x

Bruck, M., Ceci, S., Melnyk, L., & Finkelberg, D. (1999). The effect of interviewer bias

on the accuracy of children's reports and interviewer's reports. [Paper presentation] Biennial Meeting of the Society for Research in Child Development, 1999: Albuquerque, NM, United States.

Bruck, M., Melnyk, L., & Ceci, S. J. (2000). Draw it again Sam: the effect of drawing on children's suggestibility and source monitoring ability. *Journal of Experimental Child Psychology, 77*, 169–196. doi: 10.1006/jecp.1999.2560 Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419–439. doi: 10.1146/annurev.psych.50.1.419

Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review*, *22*, 520–554. doi: 10.1016/S0273-2297(02)00006-0

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology*, *18*, 947–996. doi: 10.1002/acp.1070

Bruck, M., Ceci, S. J., & Melnyk, L. (1997). External and internal sources of variation in the creation of false reports in children. *Learning and Individual Differences*, *9*, 289–316. doi: 10.1016/s1041-6080(97)90011-x

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, *38*, 217–230. doi: 10.1037/0022-3514.38.2.217

Candel, I., Memon, A., & Al-Harazi, F. (2007). Peer discussion affects children's memory reports. *Applied Cognitive Psychology*, *21*, 1191–1199. doi: 10.1002/acp.1355

Casillas, C. J., Enns, P. K., & Wohlfarth, P. C. (2010). how public opinion constrains the U.S. Supreme Court. *American Journal of Political Science, 55*, 74–88. doi: 10.1111/j.1540-5907.2010.00485.x

Ceci, S. J., & Bruck, M. (1995). Jeopardy in the courtroom: A scientific analysis of children's testimony. *American Psychological Association*. doi: 10.1037/10180-000

Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, *113*, 403–439. doi: 10.1037//0033-2909.113.3.403

Ceci, S. J., & Friedman, R. D. (2000). The suggestibility of children: Scientific research and legal implications. *Cornell Law Review*, *86*, 34–108.

Ceci, S. J., & Huffman, M. L. C. (1997). How suggestible are preschool children? Cognitive and social factors. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 948-958. doi: 10.1097/00004583-199707000-00017

Ceci, S. J., Kulkofsky, S., Klemfuss, J. Z., Sweeney, C. D., & Bruck, M. (2007). Unwarranted assumptions about children's testimonial accuracy. *Annual Review of Clinical Psychology*, *3*, 311–328. doi: annurev.clinpsy.3.022806.091354

Chaffin, M., Letourneau, E., & Silovsky, J. F. (2002). Adults, adolescents, and children who sexually abuse children: A developmental perspective. In J. E. B. Myers, L. Berliner, J. Briere, C. T. Hendrix, C. Jenny, & T. A. Reid (Eds.), *The APSAC*

*handbook on child maltreatment* (pp. 205–232). Sage Publications, Inc.

Champely, S. (2018). pwr: Basic functions for power. R package version 1.2-2. https://CRAN.R-project.org/package=pwr

Check, J. V., & Malamuth, N. M. (1983). Sex role stereotyping and reactions to depictions of stranger versus acquaintance rape. *Journal of Personality and Social Psychology*, *45*, 344–356. doi: 10.1037/0022-3514.45.2.344

*Children and teens: statistics. (n.d.).* RAINN. https://www.rainn.org/statistics/children-and-teens

Christensen, R. H. B. (2019). Ordinal–Regression models for ordinal data. R package version 2019.4-25. http://www.cran.r-project.org/package=ordinal/

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

Classen, C. C., Palesh, O. G., & Aggarwal, R. (2005). Sexual revictimization: A review of the empirical literature. *Trauma, Violence, & Abuse*, *6*, 103–129. doi: 10.1177/1524838005275087

Cohen, J. (1985). Trauma and repression. *Psychoanalytic Inquiry*, *5*, 163–189. doi: 10.1080/07351698509533580

Coker, L. S. (1990) A therapeutic recovery model for the female adult incest survivor. *Issues in Mental Health Nursing*, *11*, 109–123, doi: 10.3109/01612849009014548

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

Connolly, D. A., & Lavoie, J. A. A. (2015). Discriminating veracity between children's reports of single, repeated, and fabricated events: A critical analysis of criteria-based content analysis. *American Journal of Forensic Psychology*, *33*, 25–48.

Connolly, D. A., & Lindsay, D. S. (2001). The influence of suggestions on children's reports of a unique experience versus an instance of a repeated experience. *Applied Cognitive Psychology*, *15*, 205–223. doi: 10.1002/1099-0720(200103/04)15:2<205::AID-ACP698>3.0.CO;2-F

Connolly, D. A., & Price, H. L. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology*, *93*, 207–223. doi: 10.1016/j.jecp.2005.06.004

Connolly, D.A., & Read, J.D. (2006). Delayed prosecutions of historic child sexual abuse: analyses of 2064 Canadian criminal complaints. *Law and Human Behavior, 30*, 409-434. doi: 10.1007/s10979-006-9011-6

Connolly, D. A., Gagnon, N. C. and Lavoie, J. A. (2008[a]). The effect of a judicial declaration of competence on the perceived credibility of children and defendants. *Legal and Criminological Psychology*, *13,* 257–277. doi: 10.1348/135532507X206867

Connolly, D. A., Price, H. L., Lavoie, J. A., & Gordon, H. M. (2008[b]). Perceptions and

predictors of children's credibility of a unique event and an instance of a repeated event. *Law and Human Behaviour*, *32*, 92–112. doi: 10.1007/s10979-006- 9083-3

Connolly, D. A., Gordon, H. M., Woiwod, D. M., & Price, H. L. (2016). What children recall about a repeated event when one instance is different from the others. *Developmental Psychology*, *52*, 1038–1051. doi: 10.1037/dev0000137

Coons, P. M. (2008). *Try to remember: Psychiatry's clash over meaning, memory, and mind*. Dana Press.

Cordisco Steele, L., & National Children's Advocacy Center. (2015). Do Forensic Interview Protocols Work for Preschoolers? Huntsville, AL: National Children's Advocacy Center. Retrieved from: https://calio.org/wp-content/uploads/2019/06/do-forensic-interview-protocols-work-preschoolers.pdf

Cromer, L., & Goldsmith, R. E. (2010). Child sexual abuse myths: Attitudes, beliefs, and individual differences. *Journal of Child Abuse*, *19*, 618–647. doi: 10.1080/10538712.2010.522493

Davis, P. K., & Schwartz, G. E. (1987). Repressions and the inaccessibility of affective memories. *Journal of Personality and Social Psychology*, *52*, 155–162. doi: 10.1037//0022-3514.52.1.155

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

Desjardins T., & Scoboria A. (2007). 'You and your best friend Suzy put Slime in Ms. Smollett's desk': Producing false memories with self-relevant details. *Psychonomic Bulletin & Review*, *14*, 1090–1095. doi: 10.3758/bf03193096

DeLoache, J. S., & Smith, C. M. (1999). *Early symbolic representation*. In I. E. Sigel (Ed.), Development of mental representation: Theories and applications (pp. 61–86). Lawrence Erlbaum Associates Publishers.

Dror, I. E., Morgan, R. M., Rando, C., & Nakhaeizadeh, S. (2017). Letter to the editor – The bias snowball and the bias cascade effects: Two distinct biases that may impact forensic decision making. *Journal of Forensic Sciences*, *62*, 832–833. doi: 10.1111/1556-4029.13496

Endres, J. (1997). The suggestibility of the child witness: The role of individual differences and their assessment. *Journal of Credibility Assessment and Witness Psychology*, *1*, 44–67.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behaviour Research Methods, Instruments, & Computers*, *28*, 1–11. doi: 10.3758/BF03203630

Ernberg, E., Magnusson, M., & Landström, S. (2018). Prosecution of child sexual abuse cases involving preschool-aged children: a study of swedish cases from 2010 to 2014. *Journal of Child Sexual Abuse, 27,* 832-851. doi: 10.1080/10538712.2018.1501

Falzarano, F., & Siedlecki, K. L. (2019). Investigating the relations among different

measures of false memory. *Advances in Cognitive Psychology*, *15*, 290–300. doi: 10.5709/acp-0276-0

Farrar, M. J., & Goodman, G. S. (1992). Developmental changes in event memory. *Child Development*, *63*, 173–187. doi: 10.2307/1130911.

Feske, U. (1998), Eye movement desensitization and reprocessing treatment for posttraumatic stress disorder. *Clinical Psychology: Science and Practice, 5,* 171-181. doi: 10.1111/j.1468-2850.1998.tb00142.x

Finkelhor. D. ( 1984). *Child sexual abuse: New theory and research*. The Free Press.

Finkelhor, D., Turner, H. A., Shattuck, A., Hamby, S. L. (2013). Violence, crime, and abuse exposure in a national sample of children and youth: An update. *JAMA Pediatrics*, *167*, 614–621. doi: 10.1001/jamapediatrics.2013.42

Finnilä, K., Mahlberg, N., Santtila, P., Sandnabba, K., & Niemi, P. (2003). Validity of a test of children's suggestibility for predicting responses to two interview situations differing in their degree of suggestiveness. *Journal of Experimental Child Psychology*, *85*, 32–49. doi: 10.1016/s0022-0965(03)00025-0

Fisher, B. S., Daigle, L. E., Cullen, F. T., & Turner, M. G. (2003). Reporting sexual victimization to the police and others: Results from a national-level study of college women. *Criminal Justice and Behavior, 30*, 6–38. doi: 10.1177/0093854802239161

Flin, R., Boon, J., Knox, A., & Bull, R. (1992). Children's memories following a five-month delay. *British Journal of Psychology*, *83*, 323–336. doi: 10.1111/j.2044-8295.1992.tb02444.x

Foxman, B., Valdez, R. B., & Brook, R. H. (1986). Childhood enuresis: Prevalence, perceived impact, and prescribed treatments. *Pediatrics, 77*, 482-487.

Franklin, A., Brown, S., & Brady, G. (2018). The use of tools and checklists to assess the risk of child sexual exploitation: Lessons from UK practice. *Journal of Child Sexual Abuse*, *27*, 978-997. doi: 10.1080/10538712.2018.1510453

French, L., Sutherland, R., & Garry, M. (2006). Discussion affects memory for true and false childhood events. *Applied Cognitive Psychology*, *20*, 671–680. doi: 10.1002/acp.1219

Freud, S. (1961). *The ego and the id.* W. W. Norton & Co.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355. doi: 10.1037//0096-3445.108.3.316

Frohmann, L. (1991). Discrediting victims' allegations of sexual assault: Prosecutorial accounts of case rejections. *Social Problems*, *38*, 213–226. doi: 10.1525/sp.1991.38.2.03a00070

Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology*, *17*, 533–543. doi: 10.1002/acp.885

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining

the effects of socially encountered misinformation. *Legal and Criminological Psychology*, *9*, 215–227. doi: 10.1348/1355325041719428

Gabbert, F., Wright, D. B., Memon, A.; Skagerberg, E. M., & Jamieson, K. (2012). Memory conformity between eyewitnesses. *Court Review: The Journal of the American Judges Association. 48,* 36-43 Retrieved from: https://digitalcommons.unl.edu/ajacourtreview/382

García Marquez, G. (2014). *Living to tell the tale.* Penguin Books Ltd.

Garry, M., & Wade, K. A. (2005). Actually, a picture is worth less than 45 words: Narratives produce more false memories than photographs do. *Psychonomic Bulletin & Review*, *12*, 359–366. doi: 10.3758/bf03196385

Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S. III. (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, *83*, 347–359. doi: 10.1037/0021-9010.83.3.347

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114. doi: 10.1016/j.neuropsychologia.2013.11.010

*Glance at the career of teacher William Vahey, believed to be a prolific pedophile.* (2014, May 13). Associated Press. Retrieved March 2, 2020 from https://www.neweurope.eu/article/glance-career-teacher-william-vahey-believed-be-prolific-pedophile/

Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, *26*, 20–33. doi: 10.3758/BF03211367

Goodman, G. S., Quas, J. A., Bottoms, B. L., Qin, J., Shaver, P. R., Orcutt, H., & Shapiro, C. (1997). Children's religious knowledge: Implications for understanding satanic ritual abuse allegations. *Child Abuse & Neglect*, *21*, 1111–1130. doi: 10.1016/s0145-2134(97)00070-7

Goswick, A. E., Mullet, H. G., & Marsh, E. J., (2013). Suggestibility of stories: Can production difficulties and source monitoring explain a developmental reversal? *Journal of Cognition and Development*, 14, 607–616. doi: 10.1080/15248372.2012.710864

Gwet, K. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *The British Journal of Mathematical and Statistical Psychology*, *61*, 29–48. doi: 10.1348/000711006X126600

Haberman, C. (2014, March 9). The trial that unleashed hysteria over child abuse. *The New York Times*. https://www.nytimes.com/2014/03/10/us/the-trial-that-unleashed-hysteria-over-child-abuse.html

Hans, V. P. (2008). Jury systems around the world. *Cornell Law Faculty Publications 305.* https://scholarship.law.cornell.edu/facpub/305

Hans, V. P., & Dee, J. L. (1991). Media coverage of law: Its impact on juries and the public. *American Behavioural Scientist*, *35*, 136–149. doi: 10.1177/0002764291035002005

Havard, C. (2014). Are children less reliable at making visual identifications than adults? A

review. *Psychology, Crime & Law*, *20*, 372–388. doi: 10.1080/1068316X.2013.793334

Heaps, C. M., & Nash, M. (2001). Comparing recollective experience in true and false autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 920-930. doi: 10.1037//0278-7393.27.4.920

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

Hibbard, R. A., & Hartman, G. L. (1992). Behavioural problems in alleged sexual abuse victims. *Child Abuse & Neglect*, *16*, 755–762. doi: 10.1016/0145-2134(92)90112-5

Holliday, R. (2003). Reducing misinformation effects in children with cognitive interviews: Dissociating recollection and familiarity. *Child Development*, *74*, 728–751. doi: 10.1111/1467-8624.00565

Holliday, R. E., Brainerd, C. J., & Reyna, V. F. (2011). Developmental reversals in false memory: Now you see them, now you don't!. *Developmental Psychology*, *47*, 442–449. doi: 10.1037/a0021058

Houben, S. T. L., Otgaar, H., Roelofs, J., & Merckelbach, H. (2018). Lateral eye movements increase false memory rates. *Clinical Psychological Science, 6*, 610–616. doi: 10.1177/2167702618757658

Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child Development*, *75*, 1402–1417. doi: 10.1111/j.1467-8624.2004.00748.x

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, 60, 229–251. doi: 10.1016/j.jml.2008.10.002

Howe, M. L., Candel, I., Otgaar, H., Malone C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory*, *18*, 58–75. doi: 10.1080/09658210903476514

Howe, M. L., & Courage, M. L. (1993). On resolving the enigma of infantile amnesia. *Psychological Bulletin, 113*, 305–326. doi: 0.1037/0033-2909.113.2.305

Howe, M. L., & Wilkinson, S. (2011). Using story contexts to bias children's true and false memories. *Journal of Experimental Child Psychology,* 108, 77-95.

Howe, M. L., & Knott, L. (2015). The fallibility of memory in judicial processes: Lessons from the past and their modern consequences. *Memory, 23*, 633-656. doi: 10.10.80/09658211.2015.1010709

Hritz, A. C. (2014). Lie to me: Compliant false accusations by children (Unpublished master's thesis). Cornell University, Ithaca, NY.

Hritz, A. C., Royer, C. E., Helm, R. K., Burd, K. A., Ojeda, K., & Ceci, S. J. (2015). Children's suggestibility research: Things to know before interviewing a child. Anuario de Psicología Jurídica, 25, 3–12. doi: 10.1016/j.apj.2014.09.002

Hughes, S. (2017). American monsters: Tabloid media and the satanic panic, 1970–

2000. *Journal of American Studies, 51*, 691-719. doi: 10.1017/S0021875816001298

Huff, M. J., Bodner, G. E., & Fawcett, J. M. (2014). Effects of distinctive encoding on correct and false memory: A meta-analytic review of costs and benefits and their origins in the DRM paradigm. *Psychonomic Bulletin & Review, 22*, 349–365. doi: 3758/s13423-014-0648-8

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

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

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

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

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Joffe, H. (2011). Thematic Analysis. In D. Harper, & A. R. Thompson (Eds.), *Qualitative Research Methods in Mental Health and Psychotherapy: A Guide for Students and Practitioners*, Chichester: John Wiley & Sons, Ltd. doi: 10.1002/9781119973249.ch15

Jonker, F., & Jonker-Bakker, P. (1991). Experiences with ritualist child sexual abuse: A case study from the Netherlands. *Child Abuse & Neglect, 15,* 191–196. doi: 10.1016/0145-2134(91)90064-K

Kennedy, A. C., & Prock, K. A. (2018). 'I still feel like I am not normal': A review of the role of stigma and stigmatization among female survivors of child sexual abuse, sexual assault, and intimate partner violence. *Trauma, violence & abuse, 19*, 512–527. doi: 10.1177/1524838016673601

King, M., & Yuille, J. (1987). Suggestibility and the child witness. In S. J. Ceci, M. Toglia, & D. Ross (Eds.), *Children's eyewitness memory* (pp. 24-35). New York: Springer-Verlag. doi: 10.1007/978-1-4684-6338-5_2

Korkman, J., Juusola, A., & Santtila, P. (2014). Who made the disclosure? Recorded discussions between children and caretakers suspecting child abuse. *Psychology, Crime, & Law, 20,* 994–1004. doi: 10.1080/1068316X.2014.902455.h

Korkman, J., Santtila, P., Westeråker, M., & Sandnabba, N. K. (2008). Interviewing techniques and follow-up questions in child sexual abuse interviews. *European Journal of Developmental Psychology, 5*, 108–128. doi: 10.1080/17405620701210460

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1-26. doi: 10.18637/jss.v082.i13.

Lamb, M. E. and Brown, D. A. (2006). Conversational apprentices: Helping children

become competent informants about their own experiences. *British Journal of Developmental Psychology, 24*, 215-234. doi: 10.1348/026151005X57657

Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: a review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31*, 1201–1231. doi: 10.1016/j.chiabu.2007.03.021

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240. doi: 0.1037/0033-295X.104.2.211

Lavoie, J., Wyman, J., Crossman, A. M., & Talwar, V. (2021). Meta-analysis of the effects of two interviewing practices on children's disclosures of sensitive information: Rapport practices and question type. *Child Abuse & Neglect, 113*, 1-12. doi: 10.1016/j.chiabu.2021.104930

Lawson, M., Rodriguez-Steen, L., & London, K. (2018). A systematic review of the reliability of children's event reports after discussing experiences with a naïve, knowledgeable, or misled parent. *Developmental Review, 49*, 62–79. doi: 10.1016/j.dr.2018.06.003

Leach, A.-M., Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2004). 'Intuitive' lie detection of children's deception by law enforcement officials and university students. *Law and Human Behaviour, 28*, 661–685. doi: 10.1007/s10979-004-0793-0

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*, 568–578. doi: 10.1037/0012-1649.31.4.568

Lief, H. I., & Fetkewicz, J. (1995). Retractors of false memories: The evolution of pseudomemories. *Journal of Psychiatry & Law, 23*, 411–435. doi: 10.1177/009318539502300305

Lindberg, M. A., Keiffer, J., & Thomas, S. W. (2000). Eyewitness testimony for physical abuse as a function of personal experience, development, and focus of study. *Journal of Applied Developmental Psychology, 21*, 555–591. doi: 10.1016/s0193-3973(00)00054-x

Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology, 8*, 281-338. doi: 10.1002/acp.2350080403.

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

Loftus, E. F. & Ketcham, K. (1994). The Myth of Repressed Memory. NY: St. Martin's Press.

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

Loftus, E. F. (1997). Creating childhood memories. *Applied Cognitive Psychology, 11*, S75-S86. doi: 10.1002/(SICI)1099-0720(199712)11:7<S75::AID-ACP514>3.0.CO;2-F

Loftus, E. F. (2003). Our changeable memories: Legal and practical implications. *Nature Reviews Neuroscience,* 4, 231-234. doi: 10.1038/nrn1054

Loftus, E. F. (2004). Memories of things unseen. *Current Directions in Psychological Science, 13*, 145-147. doi: 10.1111/j.0963-7214.2004.00294.x

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361-366. doi: 10.1101/lm.94705.

London, K., & Kulkofsky, S. (2010). Factors affecting the reliability of children's reports. *New Frontiers in Applied Memory.* 119-141.

London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law, 11*, 194–226. doi: 10.1037/1076-8971.11.1.194

Maran, M. (2010). *My lie: A true story of false memory*. Chichester, United Kingdom: John Wiley and Sons Ltd.

Marx, B. P., & Sloan, D. M. (2002). The role of emotion in the psychological functioning of adult survivors of childhood sexual abuse. *Behaviour Therapy, 33*, 563-577. doi: 10.1016/S0005-7894(02)80017-X.

Mazzoni, G., & Kirsch, I. (2002). *Autobiographical memories and beliefs: A preliminary metacognitive model.* In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (p. 121–145). Cambridge University doi: 10.1017/CBO9780511489976.007

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

McGuire, K., & London, K. (2020). A retrospective approach to examining child abuse disclosure. *Child Abuse & Neglect, 99*, 104263. doi: 10.1016/j.chiabu.2019.104263

McGuire, K., London, K., & Wright, D. B. (2015). Developmental trends in false memory across adolescence and young adulthood: A comparison of DRM and memory conformity paradigms. *Applied Cognitive Psychology, 29*, 334-344. doi: 10.1002/acp.3114

Meiselman, K. C. (1990). *The Jossey-Bass social and behavioural science series. Resolving the trauma of incest: Reintegration therapy with survivors.* Jossey-Bass.

Messman-Moore T. L., Long P. J. (2003) The role of childhood sexual abuse sequelae in the sexual revictimization of women: An empirical review and theoretical reformulation. *Clinical Psychology Review, 23,* 537-71. doi: 10.1016/s0272-7358(02)00203-9.

Metzger, R.L., Warren, A.R., Shelton, J.T., Price, J., Reed, A.W., & Williams, D. (2008). Do children 'DRM' like adults? False memory production in children. *Developmental Psychology, 44,* 169181. doi: 10.1037/0012-1649.44.1.169

Millar, P. (2018, January 2nd). Why Indonesian activists are defending these school employees jailed for child sex abuse? Retrieved from: http://sea-globe.com/neil-bantleman/

Murray, L. K., Nguyen, A., & Cohen, J, A. (2015). Child sexual abuse. *Child and Adolescent Psychiatric Clinics of North America, 23,* 321-337. doi: 10.1016/j.chc.2014.01.003

Nakhaeizadeh S., Morgan R. M., Rando C., Dror I.E. (2017). Cascading bias of initial exposure to information at the crime scene to the subsequent evaluation of skeletal remains. *Journal of Forensic Sciences, 63*, 403–11. doi: 10.1111/1556-4029.13569

Nathan, D., & Snedeker, M. R. (2001). *Satan's silence: Ritual, abuse and the making of a modern American witch hunt.* iUniverse.

Newman, B. M., Lohman, B. J., & Newman, P. R. (2007). Peer group membership and sense of belonging: Their relationship to adolescent behaviour problems. *Adolescence. 42*, 241-63.

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* e2026447118. doi: 10.1073/pnas.2026447118

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*, 733–752. doi: 10.1016/s0145-2134(00)00137-x

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

Ost, J., Costall, A., & Bull, R. (2001). False confessions and false memories: A model for understanding retractors' experiences. *The Journal of Forensic Psychiatry, 12*, 549–579. doi: 10.1080/09585180127393

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

Ost, J., Granhag, P. A., Udell, J., & Roos af Hjelmsäter, E. (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. doi: 10.1080/09658210701723323

Ornstein, P. A., Merritt, K. A., Baker-Ward, L., Furtado, E., Gordon, B. N., & Principe, G. (1998). Children's knowledge, expectation, and long-term retention. *Applied Cognitive Psychology, 12*, 387–405. doi: 10.1002/(sici)1099-0720(199808)12:4<387::aid-acp574>3.0.co;2-5

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

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

Otgaar, H., Candel, I., Scoboria, A., & Merckelbach, H. (2010) Script knowledge enhances the development of children's false memories. *Acta Psychologica, 133,* 57-63. doi: 10.1016/j.actpsy.2009.09.002.

Otgaar, H., Candel, I., Merckelbach, H., & Wade, K. A. (2009). Abducted by a UFO: Prevalence information affects young children's false memories for an implausible event. *Applied Cognitive Psychology, 23*, 115–125. doi: 10.1002/acp.1445

Otgaar , H., Horselenberg, R., van Kampen, R., & Lalleman, K. (2012). Clothed and unclothed human figure drawings lead to more correct and incorrect reports of touch in children. *Psychology, Crime & Law, 18*, 641-653. doi: 10.1080/1068316X.2010.532129

Otgaar, H., Howe, M. L., Muris, P.Y, & Merckelbach, H. (2019). Associative activation as a mechanism underlying false memory formation. *Clinical Psychological Science*, *7*, 191–195. doi: 10.1177/2167702618807189.

Otgaar, H., Howe, M. L., Peters, M., Sauerland, M., & Raymaekers, L. (2013). Developmental trends in different types of spontaneous false memories: Implications for the legal field. *Behavioural Sciences and the Law,* 31, 666–682. doi: 10.1002/bsl.2076

Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology: General, 145,* 31-55. doi: 10.1037/xge0000127

Otgaar, H., de Ruiter, C., Howe, M. L., Hoetmer, L., & van Reekum, P. (2017). A case concerning children's false memories of abuse: Recommendations regarding expert witness work. *Psychiatry, Psychology and Law, 24*, 365-378. doi: 10.1080/13218719.2016.1230924

Otgaar, H., Howe, M. L., Merckelbach, H., & Muris, P. (2018). Who is the better eyewitness? Sometimes adults but at other times children. *Current Directions in Psychological Science, 27*, 378–385. doi: 10.1177/0963721418770998

Otgaar, H., Howe, M. L., Patihis, L., Merckelbach, H., Lynn, S. J., Lilienfeld, S. O., & Loftus, E. F. (2019). The return of the repressed: The persistent and problematic claims of long-forgotten trauma. *Perspectives on Psychological Science, 14*, 1072–1095. doi: 10.1177/1745691619862306

Otgaar, H., La Rooy, D., Horselenberg, R., Herskowitz, I., De Ruiter, C., Blezer, L. Kidane, R., & Kollau, R. (2019). Assessing the quality of child investigative interviewing in the Netherlands. *Applied Cognitive Psychology, 23,* 889-897. doi:

10.1002/acp.3521.

Padock, R. C. *Jakarta International School sexual-abuse investigation widens* (2014, July 12). The Wall Stree Journal. https://www.wsj.com/articles/jakarta-international-school-sexual-abuse-investigation-widens-1405145862

Patihis, L., Frenda, S., J., LePort, A., K., R., Petersen, N., Nichols, R., M., Stark, C. E. L., …, & Loftus, E. F. (2013). False memories in highly superior autobiographical memory individuals. *Proceedings of the National Academy of Sciences, 110*, 20947-20952. doi: 10.1073/pnas.1314373110

Patihis, L., Frenda, S., J., & Loftus, E. F. (2018). False memory tasks do not reliably predict other false memories. *Psychology of Consciousness: Theory, Research and Practice, 5*, 140-160. doi: 10.1037/cns0000147

Pezdek, K., Finger, K., & Hodge, D. (1997). Planting false childhood memories: The role of event plausibility. *Psychological Science, 8*, 437-441. doi: 10.1111/j.1467-9280.1997.tb00457.x.

Pezdek, K., Blandon-Gitlin, I., Lam, S., Ellis Hart, R., & Schooler, J. W. (2006) Is knowing believing? The role of event plausibility and background knowledge in planting false beliefs about the personal past. *Memory & Cognition 34***,** 1628–1635. doi: 10.3758/BF03195925.

Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., … Brodie, L. (2004). Detecting deception in children: Event familiarity affects Criterion-Based Content Analysis ratings. *Journal of Applied Psychology, 89*, 119–126. doi: 10.1037/0021-9010.89.1.119

Pipe, M. E., Lamb, M. E., Orbach, Y., & Esplin, P. W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review, 24*, 440–468. doi: 10.1016/j.dr.2004.08.006

Poole, D. A., & Lamb, M. E. (1998). Investigative interviews of children: A guide for helping professionals. *American Psychological Association*. doi.org:

Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied*, *7*, 27–50. doi: 10.1037/1076-898X.7.1.27

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27*, 975–986. doi: 10.1037/0012-1649.27.6.975

Poole, D. A. & Wolfe, M. A. (2009). Child development: Normative sexual and nonsexual behaviours that may be confused with symptoms of sexual abuse. *The Evaluation of Child Sexual Abuse Allegations: A Comprehensive Guide to Assessment and Testimony,* 101-128.

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

10.1023/a:1022344128649.

Price, H. L., & Connolly, D. A. (2004). Event frequency and children's suggestibility: A study of cued recall responses. *Applied Cognitive Psychology, 18*, 809–821. doi: 10.1002/acp.1059

Price, H. L., Connolly, D. A., & Gordon, H. M. (2006). Children's memory for complex autobiographical events: Does spacing of repeated instances matter? *Memory, 14*, 977–989. doi: 10.1080/09658210601009005

Principe, G. F., & Ceci, S. J. (2002). "I saw it with my own ears": the effects of peer conversations on preschoolers' reports of nonexperienced events. *Journal of experimental child psychology, 83*, 1–25. doi: 1016/s0022-0965(02)00120-0

Principe, G. F., Cherson, M., DiPuppo, J., & Schindewolf, E. (2012). Children's natural conversations following exposure to a rumor: Linkages to later false reports. *Journal of Experimental Child Psychology, 113*, 383–400. doi: 10.1016/j.jecp.2012.06.006

Principe, G. F., Dipuppo, J., & Gammel, J. (2013). Effects of mothers' conversation style and receipt of misinformation on children's event reports. *Cognitive Development, 28,* 260-271. doi: 10.1016/j.cogdev.2013.01.012

Principe, G. F., & Schindewolf, E. (2012). Natural conversations as a source of false memories in children: Implications for the testimony of young witnesses. *Developmental Review: DR, 32*, 205–223. doi: 10.1016/j.dr.2012.06.003

Principe, G. F., Guiliano, S., & Root, C. (2008). Rumor mongering and remembering: How rumors originating in children's inferences can affect memory. *Journal of Experimental Child Psychology, 99*, 135–155. doi: 10.1016/j.jecp.2007.10.009

Principe, G. F., & Smith, E. (2007). The tooth, the whole tooth and nothing but the tooth: How belief in the Tooth Fairy can engender false memories. *Applied Cognitive Psychology, 22,* 625-642. doi: 10.1002/acp.1402

Processor Éticos. (n.d.). Conselho Federal de Psicologia. https://site.cfp.org.br/servicos/orientacao-e-etica/processos-eticos/

Putnam F. W. (2003). Ten-year research update review: child sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 269–278. doi: 10.1097/00004583-200303000-00006

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

R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org

RAINN. (2019). Victims of sexual violence: Statistics. Retrieved August 1, 2019, from https://www.rainn.org/statistics/victims-sexual-violence

Randall, M. (2010). Sexual assault law, credibility, and 'ideal victims': Consent, resistance, and victim blaming. *Canadian Journal of Women and the Law, 22*, 397-433. doi: 10.3138/cjwl.22.2.397

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1–75. doi: 10.1016/1041-6080(95)90031-4

Roberts, K. P., Lamb, M. E., & Sternberg, K. J. (2004). The effects of rapport-building style on children's reports of a staged event. *Applied Cognitive Psychology, 18*, 189–202. doi: org/10.1002/acp.957

Roediger, H. L. (2001). Psychology of reconstructive memory. In Smelser, N. J. & Baltes, P. B (Ed), *International Encyclopedia of the Social & Behavioural Sciences* (pp. 12844-12849). Pergamon.

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

Roediger, H. L., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*, 365–371. doi: 10.3758/bf03196174

Roediger, H. L. III, Zaromb, F. M. & Butler, A. C. (2009). The role of repeated retrieval in shaping collective memory. *Memory in Mind and Culture*, 138-170. doi: 10.1017/CBO9780511626999.00

Rosenthal, R. (1995). State of New Jersey v. Margaret Kelly Michaels: An overview. *Psychology, Public Policy, and Law, 1*, 246-271. doi: 10.1037/1076-8971.1.2.246

Rossen, B. (1992). Response to the Oude Pekela incident and the accusations of Drs. F. Jonker and I. Jonker-Bakker. *Journal of Psychology and Theology, 20*, 263-266. doi: 10.1177/009164719202000321

Rowell, L. L, Polush E. Y., Riel M., & Bruewer, A. (2015) Action researchers' perspectives about the distinguishing characteristics of action research: a Delphi and learning circles mixed-methods study, *Educational Action Research, 23*, 243-270, doi: 10.1080/09650792.2014.990987

Rubínová, E., Blank, H., Ost, J., & Fitzgerald, R. (2020). Structured word-lists as a model of basic schemata: deviations from content and order in a repeated event paradigm**.** *Memory, 28,* 309-322. doi: 10.1080/09658211.2020.1712421.

Saint-Martin, P., Bouyssy, M., & O'Byrne, P. (2007). Analysis of 756 cases of sexual assault in Tours (France): medico-legal findings and judicial outcomes. *Medicine, Science and the Law, 47*, 315–324. doi: 10.1258/rsmmsl.47.4.315

Sauerland, M., & Otgaar, H. (2021). Teaching psychology students to change (or correct) controversial beliefs about memory works, *Memory*, doi: 10.1080/09658211.2021.1874994

Sauerland, M., Brackmann, N., & Otgaar, H. 2018). Rapport: Little effect on children's, adolescents', and adults' statement quantity, accuracy, and suggestibility. *Journal of Child Custody, 15*¸268-285. doi: 10.1080/15379418.2018.1509759

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures.* Lawrence Erlbaum.

Schreiber, N., Wentura, D., & Bilsky, W. (2001). 'What else could he have done?'

Creating false answers in child witnesses by inviting speculation. *Journal of Applied Psychology, 86*, 525–532. doi: 10.1037/0021-9010.86.3.525

Schreiber, N., Bellah, L. D., Martinez, Y., McLaurin, K. A., Strok, R., Garven, S., & Wood, J. M. (2006). Suggestive interviewing in the McMartin Preschool and Kelly Michaels day care abuse cases: A case study. *Social Influence, 1*, 16-47. doi: 10.1080/15534510500361739

Schönbucher, V., Maier, T., Mohler-Kuo, M., Schnyder, U., Landolt, M.A. (2012). Disclosure of child sexual abuse by adolescents: a qualitative in-depth study. *Journal of Interpersonal Violence. 27*, 3486-3513. doi: 10.1177/0886260512445380

Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology. 18,* 79-07. doi: 10.1002/acp.1062.

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

Sgroi, S. (1982). *Handbook of Clinical Intervention in Child Sexual Abuse* (illustrated, reprint ed.). Simon and Schuster.

Shaw, J., & Porter, S. (2015). Constructing rich false memories of committing a crime. *Psychological Science, 26*, 291-301. doi: 10.1177/0956797614562862.

Shaw, J., & Vredeveldt, A. (2019). The recovered memory debate continues in Europe: Evidence from the United Kingdom, the Netherlands, France, and Germany. Clinical Psychological Science, 7, 27–28. doi: 10.1177/2167702618803649

Shead, K. (2014). Responding to historical child sexual abuse: A prosecution perspective on current challenges and future directions. *Current Issues in Criminal Justice, 26*, 55-73. doi: 10.1080/10345329.2014.12036007

Short, J. L., & Bodner, G. E. (2011). Differentiating accounts of actual, suggested and fabricated childhood events using the judgment of memory characteristics questionnaire. *Applied Cognitive Psychology, 25*, 775-781. doi: 10.1002/acp.1756.

Sjöberg, R. L. (1997). False allegations of satanic abuse: Case studies from the witch panic in Rättvik 1670-71. *European Child & Adolescent Psychiatry, 6*, 219–226. doi: 10.1007/bf00539929

Slackman, E., & Nelson, K. (1984). Acquisition of an unfamiliar script in story form by young children. *Child Development*, *55*, 329 - 340.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.

Spohn, C., Beichner, D., & Frenzel, E. (2001). Prosecutorial justifications for sexual assault case rejection: Guarding the 'gateway to justice'. *Social Problems, 48*. 206-235. 10.1525/sp.2001.48.2.206.

Stewart, N. D. (2011). An experience reducing toilet flushing noise reaching

adjacent offices. *The Journal of Acoustical Society of America, 129,* 2605-2605. doi: 10.1121/1.3588639

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

Strange, D., Sutherland, R., & Garry, M. (2006). Event plausibility does not determine children's false memories. *Memory, 14*, 937–951. doi: 10.1080/09658210600896105

Sumampouw, N. E. J., Otgaar, H., La Rooy, D., & De Ruiter, C. (2019). The quality of forensic child interviewing in child sexual abuse cases in Indonesia. *Journal of Police and Criminal Psychology, 35*, 170–181. doi: 10.1007/s11896-019-09342-5

Sumampouw, N. E. J., de Ruiter, C., & Otgaar, H. (2021). Potential for police investigator bias: the impact of child sexual abuse victims' background characteristics on perceived statement credibility, case outcome and quality of interview questions. *Police Practice and Research, 1 – 18.* doi: 10.1080/15614263.2021.1952

Summit, R., & Kryso, J. (1978). Sexual abuse of children: A clinical spectrum. *American Journal of Orthopsychiatry, 48,* 237–251. doi: 10.1111/j.1939-0025.1978.tb01312.x

Sutherland, R., & Hayne, H. (2001). The effect, of post event information on adult's eyewitness reports. *Applied Cognitive Psychology, 15*, 249-263. doi: 10.1002/acp.700

Talmi, D., Luk, B. T. C., McGarry, L. M., Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language, 56*, 555-574. doi: 10.1016/j.jml.2007.01.002

Tenney, E. R., Spellman, B. A., & MacCoun, R. J. (2008). The benefits of knowing what you know (and what you don't): How calibration affects credibility. *Journal of Experimental Social Psychology, 44*, 1368-1375. doi: 10.1016/j.jesp.2008.04.006

The National Center for Victims of Crime. (2011). Statistics on perpetrators of child sexual abuse. Retrieved August 1, 2019, from https://victimsofcrime.org/media/reporting-on-child-sexual-abuse/statistics-on-perpetrators-of-csa

Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do? Suggestive interviewing and the accuracy of children's accounts. *Law and Human Behaviour, 21*, 405–426. doi: 10.1023/A:1024859219764

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*, 1–12. doi: 10.1037/h0080017

Van Dyke, R. M., and S. E. Alcock. 2003. *'Archaeologies of Memory: An Introduction.*' In Archaeologies of Memory, edited by R. M. Van Dyke and S. E. Alcock, 1–14. Oxford: Blackwell.

van Koppen, P. J. and Mackor, A. R. (2019). A Scenario Approach to the Simonshaven Case. *Topics in Cognitive Science*, 1-20. doi: 10.1111/tops.12429

van Veldhuizen, T. S., Horselenberg, R., Landström, S., Granhag, P. A., & van Koppen, P. J. (2017). Interviewing asylum seekers: A vignette study on the questions asked to assess credibility of claims about origin and persecution. *Journal of Investigative*

*Psychology and Offender Profiling, 14*, 3-22. doi: 10.1002/jip.1472

Wade, K. A., Garry, M., Nash, R. A., & Harper, D. N. (2010). Anchoring effects in the development of false childhood memories. *Psychonomic Bulletin & Review, 17,* 66-72. doi: 10.3758/pbr.17.1.66.

Wade, K. A., Garry, M., Read, J. D., & Lindsay, D. S. (2002). A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin & Review, 9*, 597-603. doi: 10.3758/bf03196318.

Wade, K. A., Garry, M., & Pezdek, K. (2018). Deconstructing rich false memories of committing a crime: Commentary on Shaw and Porter (2015). *Psychological Science, 29*, 471-476. doi: 10.1177/0956797617703667.

Walker, H. E., Freud, J. S., Ellis, R. A., Fraine, S. M., & Wilson, L. C. (2019). The prevalence of sexual revictimization. *Trauma, Violence, & Abuse, 20,* 67-80. doi: 10.1177/1524838017692364

Warren, A. R., & McGough, L. S. (1996). Research on children's suggestibility. *Criminal Justice and Behaviour, 23*, 269–303. doi: 10.1177/0093854896023002003

Weinsheimer, C. C., Coburn, P. I., Chong, K., MacLean, C. L., & Connolly, D. A. (2017). Perceptions of credibility for a memory report of a single versus repeated event. *Applied Cognitive Psychology, 31*, 414-423. doi: 10.1002/acp.3340

Wilde, Oscar. 1990. *The Importance of Being Earnest*. Dover Thrift Editions. New York, NY: Dover Publications.

Wolfe, D. A., Sas, L., & Wekerke. (1994) Factors associated with the development of posttraumatic stress disorder among child victims of sexual abuse. *Child Abuse & Neglect, 18*, 37-50. doi: 10.1016/0145-2134(94)90094-9.

Wolfe, V. V. (2007). *Child sexual abuse*. In E. J. Mash & R. A. Barkley (Eds.), Assessment of childhood disorders (pp. 685–748). The Guilford Press.

Wright, D. B., Self, G., Justice, C. (2000). Memory conformity: exploring misinformation effects when presented by another person. *British Journal of Psychology, 91*, 189-202. doi: 10.1348/000712600161781

Wright, D. B., & Schwartz, S. L. (2010). Conformity effects in memory for actions. *Memory and Cognition, 38*, 1077-1086. doi: 10.3758/MC.38.8.107

Wyatt, W. J. (2002). What was under the McMartin preschool? A review and behavioural analysis of the 'tunnels' find. *Behaviour and Social Issues, 12*, 29–39. doi: 10.5210/bsi.v12i1.77

Young, M. (2008). The day care ritual abuse moral panic: A sociological analysis. *Sociology Compass, 2,* 1719-1733. doi: 10.1111/j.1751-9020.2008.00169.x.

Yuhas, A. (2021, March 31). *It's Time to Revisit the Satanic Panic*. The New York Times. https://www.nytimes.com/2021/03/31/us/satanic-panic.html

Zaragoza, M. S., Michell, K. J. (1996). Repeated exposure to suggestion and the creation of false memories. *Psychological Science, 7*, 294-300.

Zaragoza, M. S., Mitchell, K. J., Payment, K., & Drivdahl, S. (2011). False memories

for suggestions: The impact of conceptual elaboration. *Journal of Memory and Language, 64,* 18–31. doi: 10.1016/j.jml.2010.09.004

Frost, P., Lacroix, D., & Sanborn, N. (2003). increasing false recognition rates with confirmatory feedback: A phenomenological analysis. *The American Journal of Psychology*, *116*, 515–525. doi: 10.2307/1423658

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

## Impact

### *Main Goal of this Thesis*

The main objective of the current thesis was to examine different themes within cases of child sexual abuse revictimization. Specifically, this thesis started with the Jakarta International School (JIS) case, which features accusations of repeated child sexual abuse against two schoolteachers and five outsourced cleaners at three kindergartens. The JIS case shares commonalities with so-called day-care abuse cases, such as the McMartin preschool and Kelly Michaels case. Such cases often present several features that make the children's claims seem a product of high exposure to external suggestion, potentially resulting in false allegations and false memories.

In the JIS case the accused were convicted solely based on children's statements, as no corroborative evidence was found to substantiate children's accusations. Despite their rarity, when such cases occur, they can render significant harm to families involved, while also taxing investigations by the legal system. The JIS case is a relevant and recent illustration of the complexity of themes that can arise in such cases. For instance, a case like the JIS in which revictimisation is present may involve potential false memory formation and credibility assessment issues. In the current thesis, I aimed to investigate false memory formation in different contexts and credibility assessment for revictimised victims of sexual abuse.

### *Main Findings*

Analysing the JIS case in light of relevant research, I concluded that the techniques employed to interview and guide the children through the process of investigation were not in accordance with the guidelines of interviewing protocols recognised by the scientific community as good practice. The main discrepancies between such interviewing guidelines and the techniques used to interview the children in the JIS case concerns the use of suggestive techniques. For instance, children were hardly ever invited to explain what happened in their own words in an open manner. Mostly, children were asked from the outset what had happened during the supposed occasions of abuse. That is, interviewers already assumed some kind of abuse had taken place without necessarily hearing this from the children.

In the JIS case, the suspicions of abuse came from the parents of the children involved before the children had disclosed anything about being abused in school. In this and similar cases, it would be essential for the investigation to listen to the children's version of what happened. An important feature of the JIS case, which corroborates the conclusion that children's claims were rendered unreliable, is the evolution of the allegations. That is, in the first months of the investigation of the case children vehemently denied being sexually abused in school. However, during the course of successive suggestive interviews their statements changed. Interestingly, such issues did

not seem to affect the perceived credibility of the children's statements in the eyes of the professionals involved in the investigation and the trial of the case.

The issues observable in the JIS case and other dubious day-care abuse cases, made me decide to focus my research project on both credibility assessment for sexual abuse victims and false memory formation in different contexts. In the study in which I looked into credibility assessment in sexual abuse victims, I observed that professionals in the legal system viewed the statements of victims who had been sexually abused more than once less credible than the statements of victims who had experienced one incident of sexual abuse. Furthermore, the perceived credibility of the statements by professionals in the legal system did not differ between child or adult victims.

Concerning false memory formation, I examined the effects of suggestion on children's memory statements in a method where children receive suggestion about an event they witnessed and a method in which children did not receive suggestion and were allowed to freely recall elements they had been asked to memorise. Following similar results in the literature, this project showed that children do produce reliable memory statements when asked to freely recall what they remember from a given event. Moreover, this project showed that children are highly prone to external suggestion, and hence may produce unreliable memory statements under those circumstances. Considering that many dubious cases of sexual abuse that come to the attention of the legal system contain allegations of revictimisation, in my last study, I investigated false memory formation for repeated events. In this novel project, I found that it is not more difficult to form false memories for repeated events compared to a single event.

***Scientific Relevance***

My thesis offers useful information to the professionals in the legal field in terms of addressing the (perhaps) rare but pressing matter of potential false memories of repeated events of sexual abuse. Similar to previous case reports of potential false memory formation or false allegations, the JIS case study is a reminder to researchers that our work investigating this issue is still highly relevant in the present era. Furthermore, this case study is an example of the dangers of carrying out an investigation using suggestive questioning. Specifically, the field of false memory implantation research has been criticised for its lack of ecological validity (Blizard & Shaw, 2019; Brewin & Anderson, 2017). Because the corpus of previous research in false memory implantation for autobiographical events investigated memory for a single event, some critics doubted if such research would be applicable to real world cases. Potential cases of false memory formation for autobiographical events often concern an event that has happened more than once.

By demonstrating that false memories of repeated events can be created in laboratory conditions with the same ease as false memories of single events, I have addressed an important critique concerning lack of ecological validity in false memory

research (see Otgaar et al., in press). This is important because my work shows that false memory implantation findings can be generalised to legal cases concerning false memories. Furthermore, this finding opens the door to new directions within the false memory implantation field. For instance, based on the findings in this thesis, Otgaar and colleagues (in press) replicated my findings showing that false memories for repeated events can be elicited in laboratory conditions. This shows the reliability of my work thereby adding to the generalisability of false memory research. Alongside the JIS case, my false memory implantation of repeated events study shows the negative effect of suggestion on memory, stressing how its consequences may severely affect communities and families, as well as burden the legal system.

***Target Group and Dissemination***

Besides the scientific community, the main target group of the work performed for the current thesis are practitioners working in the legal field. Every chapter in this thesis corresponds to a project with direct application to the praxis of professionals performing child interviewing, statement credibility assessment, and child sexual abuse cases. Such professionals may work at the police station, child protection services, in the court system as well as at a variety of non-governmental organisations. Moreover, these professionals may have different training backgrounds. For example, in South American countries it is common for child sexual abuse victims to be interviewed and assisted by clinical or forensic psychologists who work for the court. Regardless of their background, the work that composes the current thesis was mainly aimed at the group of professionals around the globe who can benefit from novel findings that are relevant for child interviewing.

I have shown the harm that suggestive interviewing techniques can do with children's memories. Hence, my findings underscore the importance of evidence-based practices when interviewing children. Such evidence-based practice within this branch of the legal field has been the target of a plethora of scientific activities, including peer-reviewed publications, guidelines and protocols. For example, target groups that might profit from my findings could be child interviewers working at for instance the police or child protection. My work shows the danger of using suggestive questions and that empirically validated interview protocols should be used to minimize the use of suggestion. For example, the National Institute of Child Health and Human Development (NICHD; Lamb et al., 2007) protocol developed for child interviewing is based on scientific findings about child suggestibility, memory, communication skills, social understanding, and social tendencies. This protocol has shown to be effective in reducing interviewer's reliance on the use of suggestive questions. The work presented in the current thesis stress the importance of evidence based child interviewing trainings for relevant practitioners to reduce the use of suggestive questions during child interviews.

In terms of dissemination, Chapter V in this thesis has been presented at EAPL (European Association of Psychology and Law) and SARMAC (Society for Applied

Research in Memory and Cognition), international conferences that attract both academics and practitioners. Chapter I and V have been presented at TU Dortmund in a colloquium aimed at students and practitioners. Chapters I, III and V are published in open access journals. More importantly, the work presented in Chapter I was published in a journal that publishes articles for a general academic audience. Publishing in such journals is important as they have the potential to foment dialogue between different academics and to reach interested parties besides psychologists, such as legal professionals.

Furthermore, to increase the reach of my findings to a diverse group of professionals, this work could also be posted on blogs (e.g. Psychology Today), presented at other practitioners' conferences (e.g. International Investigative Interview Research Group Conference, Congresso Internacional de Psicologia Jurídica e Direito Penal, Rapid Fire Conference on Investigative Interviewing, Congresso Internacional de Jornalismo Investigativo), and be included in workshops and symposia. Given the public influence on a variety of legal matters, there is a pressing need for disseminating scientific findings not only to pertinent professionals but also to the general audience. Hence, it may be relevant to broadcast the findings in the current thesis in social media, via written posts, videos and podcasts (e.g. available on podcast platforms: False Memory Deutschland, True Crime False Memories, and Bad People available at https://www.bbc.co.uk/programmes/p08lj2sz/episodes/downloads) fomenting the general understanding about the topic and bridging the gap between academia and society.

## Summary

Revictimisation is common in child sexual abuse cases. That is, if a child has suffered from any form of sexual abuse, there is a high likelihood that the same child will experience such abuse again. Only a small proportion of sexual abuse cases is reported to the legal authorities. When the victim is a child, it becomes even more challenging to prosecute sexual abuse cases. Children's lack of independence from adults makes it difficult for them to identify that a crime is taking place, and consequently, it is difficult for them to report the crime to the police. After entering the legal system, statements of child victims are often evaluated in terms of their credibility and reliability. Sometimes, these evaluations conclude that children's statements concerning abuse are based on false memories.

In the 1980s and 1990s, a waterfall of child sexual abuse revictimisation cases were reported. Some of these cases included signs of suggestive interviewing tactics that might have fostered false reports in children. Cases in which an alleged child victim falsely remembers being sexually abused several times during their childhood likely do not represent most of the child sexual abuse cases that are reported to the legal system. However, cases of false memory for child sexual abuse have proven to be damaging to the legal system for they are likely to consume as much time and as many financial resources as any sexual abuse case. In face of the wave of dubious child sexual abuse cases in the 1980s and the 1990s, such as the McMartin Preschool case, researchers started to explore topics such as child suggestibility and false memory formation. A rich body of literature indicates that given the opportunity to freely report their memories, children can provide reliable statements. Furthermore, several studies have demonstrated that researchers have been able to implant an entire false autobiographical event in participants' (both adults and children) memory in laboratory conditions. Because of this lack of focus given to the implantation of *repeated* autobiographical events, research on false memory implantation has received criticism in terms of its applicability to legal cases.

To put the work conducted in the current thesis into context, I present in Chapter 2 the Jakarta International School (JIS) case, which took place in 2014 in Indonesia. In the JIS case, parents of three kindergartens reported to the police that their children had been sexually abused several times in the course of six months within the school facilities. Apart from the presence of parents' and children's statements, the prosecution of the case lacked corroborating evidence against the accused. Based on official and unofficial reports collected by the defence in the case, we analysed the JIS case in the format of a case report, while consulting the current research literature about the major issues aligned to the topics of the current thesis. Chapters 3, 4 and 5 present studies on three relevant topics for the JIS and similar cases: child suggestibility and false memory formation within a co-witness context (Chapter 3), credibility attested to revictimised sexual abuse victims (Chapter 4), and implantation of false memories for repeated events

of child sexual abuse (Chapter 5).

In Chapter 2, I concluded that the use of interviewing techniques understood as suggestive by the current body of research on child interviewing jeopardised the integrity of children's reports and consequently the investigation in the JIS case. Therefore, in Chapter 3, I used a memory conformity paradigm to investigate children's propensity to have their memory contaminated by the report of a co-witness. The main take of Chapter 3 is that children are more likely to form false memories when influenced by their peers' memory report compared to older developmental groups. This finding contrasts with the development reversal effect observed when children are invited to freely report what they remember from a given event. That is, under those conditions, children outperform older developmental groups in the DRM word learning task, making fewer memory mistakes. Another issue related to the JIS and similar cases concerns the credibility ascribed to sexual abuse victims who report victimisation.

In Chapter 4, I showed that legal system practitioners perceived reports containing sexual abuse revictimisation as less credible compared to reports containing one occurrence of sexual abuse. Despite supporting previous work on credibility assessment of memory reports for repeated events (unrelated to sexual abuse), our results do not align with what happened in the JIS case. That is, all those involved in the investigation and prosecution of the JIS case did not deem children's reports of repeated sexual abuse had low credibility. Finally, when looking into the JIS and similar cases, after repeated exposure to suggestion, false reports are a likely outcome that can occur due to false memory formation.

Similar to the JIS case, other cases also feature children reporting dubious memories of child sexual abuse revictimisation. As much as the research literature on false memory implantation had covered several aspects that could play a role in such process, we lacked a study demonstrating that false memories for repeated events could be elicited in laboratory conditions. In Chapter 5, I addressed relevant criticism of false memory implantation studies, showing that false memories for repeated events are not more difficult to implant compared to false memories for a single event. The research reported in the current thesis is relevant to the practice of investigation in child sexual abuse revictimisation cases and potential false memory formation.

## Sammanfattning På Svenska

Sexuella övergrepp på barn karaktäriseras ofta av reviktimisering. Det vill säga: ett barn som har utsatts för någon form av sexuellt våld löper hög risk att drabbas av sådant våld minst en gång till. Tyvärr anmäls endast en liten andel av de sexuella övergreppen till rättssystemet. I fall med sexuella övergrepp är det ännu svårare att väcka åtal när offret är ett barn. En anledning till detta är att barn är beroende av vuxna för att kunna tolka samhällets normer, vilket gör det svårt för dem både att inse att ett brott håller på att begås och att anmäla brottet till polisen. Vid kontakt med rättssystemet görs ofta en bedömning av trovärdigheten i de påståenden som offren för sexuella övergrepp för fram. Ibland upptäcker man att anmälningar om övergrepp på barn inte grundas på faktiska händelser utan på falska minnen.

Under 1980- och 1990-talen uppstod en flodvåg av anmälningar om reviktimiseringsfall, kopplade till sexuella övergrepp mot barn. I en del av dessa fall fanns det indikationer på suggestiva intervjutekniker som sannolikt ledde till att barn avgav falska uppgifter. Fall där ett påstått offer inkorrekt uppger sig ha minnen av att i späd ålder ha utsatts för ett flertal sexuella övergrepp utgör troligen inte merparten av de sexuella övergrepp mot barn som anmäls till rättssystemet. Men fall med falska minnen av sexuella övergrepp mot barn har visat sig vara skadliga för rättssystemet, eftersom de rimligen kräver lika mycket tid och lika stora ekonomiska resurser som vilket annat fall av sexuella övergrepp som helst. I svallvågorna efter tvivelaktiga rapporter om sexuella övergrepp mot barn under 1980- och 1990-talen, som McMartin Preschool-fallet, stärktes forskningslitteraturen avsevärt i ämnet barns suggestibilitet och bildandet av falska minnen. Litteraturen har till exempel visat att barn kan göra pålitliga utsagor om de bara får möjlighet att fritt berätta sina minnen. Men flera studier har också visat att helt falska självbiografiska händelser kan planteras i studiedeltagares minnen (hos både vuxna och barn). Men forskning på plantering av falska minnen har fått relevant kritik med avseende på deras applicerbarhet på rättsfall, eftersom forskningen ännu inte har studerat plantering av falska händelser som uppges ha inträffat upprepade gånger.

I denna avhandling används olika metoder för att diskutera och undersöka särdrag hos möjliga fall av falska minnen av reviktimisering efter sexuella övergrepp mot barn, inklusive plantering av falska minnen av upprepade självbiografiska händelser. I kapitel 2 beskriver jag specifikt fallet som utspelade sig i Indonesien 2014, vid Jakarta International School (JIS). I JIS-fallet anmälde föräldrarna till tre dagskolebarn till polisen att deras barn, vid flera tillfällen under loppet av sex månader, hade utsatts för sexuella övergrepp i skolans lokaler. Utöver föräldrarnas och barnens uppgifter saknade kärandesidan i fallet styrkande bevis mot de åtalade. Utifrån officiella och icke-officiella uppgifter från svarandesidan analyserades JIS-fallet i form av en fallrapport i vilken slutsatsen blev att barnens vittnesmål om upprepade övergrepp sannolikt var falska.

I kapitel 3, 4 och 5 fördjupas diskussionen kring falska minnen av reviktimisering

i tre relevanta ämnen:

- barns suggestibilitet och bildande av falska minnen i ett sammanhang av samtidiga vittnen
- trovärdigheten hos påståenden om sexuella övergrepp med reviktimisering
- plantering av falska minnen av upprepade tillfällen av sexuella övergrepp på barn.

I kapitel 3 beskrivs specifikt en studie av paradigmet minneskonformitet som undersöker barns benägenhet att låta minnet påverkas av utsagor av ett annat vittne. Huvudpoängen i kapitel 3 är att barn har en större benägenhet att bilda falska minnen genom påverkan från ett jämnårigt barns minnesuppgifter än av ett äldre barns uppgifter. Detta fynd kontrasterar mot omvända utvecklingseffekten, där yngre barn presterar bättre än äldre barn i en uppgift med falska minnen som lockar fram spontana falska minnen genom igenkänning enligt DRM-paradigmet och via fri återgivning. Vid fri återgivning ur minnet gör yngre barn färre misstag än äldre barn.

Ytterligare en fråga, förknippad med JIS och liknande fall, gäller trovärdigheten hos offer för sexuella övergrepp som anmäler reviktimisering. I kapitel 4 återges hur jurister som praktiserar inom rättssystemet granskade två skildringar som innehöll anmälningar om sexuella övergrepp. En skildring beskrev reviktimisering medan den andra beskrev ett tillfälle av sexuellt övergrepp. De praktiserande juristerna uppfattade anmälningar om reviktimisering av sexuellt övergrepp som mindre trovärdiga än anmälningar som enbart innehöll en övergreppshändelse. Trots tidigare stödjande arbeten på trovärdighetsvärdering av minnesuppgifter om upprepade händelser (utan koppling till sexuella övergrepp), stämmer våra resultat inte överens med det som hände i JIS-fallet. Det vill säga, ingen av de som var inblandade i utredning och åtal i JIS-fallet bedömde att barnens uppgifter om upprepade sexuella övergrepp hade låg trovärdighet. Slutligen, i samband med JIS och liknande fall är falska anmälningar ett sannolikt inslag, vilka kan uppstå på grund av falska minnen uppkomna genom lång exponering.

Till sist, i kapitel 5 avhandlar jag relevant kritik mot litteraturen om plantering av falska minnen. Jag jämförde plantering av falska minnen av enskilda respektive upprepade händelser genom att anpassa planteringsparadigmet till plantering av upprepade händelser. Resultaten, såväl deltagarnas egenskattningsformulär av minne och uppfattning och intervjuarnas skattning av deltagarnas minne av och uppfattning om den falska händelsen, visade att det inte är svårare att plantera falska minnen av upprepade händelser än falska minnen av en enskild händelse. Sammantaget bidrar arbetet i denna avhandling till forskningen kring de praktiska frågorna av reviktimisering vid sexuella övergrepp och eventuell bildning av falska minnen.

## Acknowledgements

The date was the 9th of April of 2020. As my face warmed up under the candid light beams that formed the golden hour, I closed my eyes and tried to ground myself. I needed to write this very thesis, but my thoughts were nebulous. We were deep within the lockdown due to the Covid-19 pandemic. I was spending most of my time completely alone in the 1714 home that had sheltered my varied housemates, including my paranymph Carey, and myself for four years. The flowers I had planted in the previous November were coming through, adding encouraging colours to the predominantly grey days. I had a safe home, loving people in my life, a purpose, a thesis to write, and yet, I felt depleted of impulse. Whilst the inspiration was fully absent, the urge to close a cycle was more alarming than it had ever been. Sadly, when I recall this day, it was just the beginning of an extraordinary period in my life, that I shall never forget even though I cannot quite recall it properly. The days were in essence the same, and hence there is very little that I can actually think of. However, I do remember the 9th of April 2020. I am not sure if it was the sun light on my closed eyelids (I always liked the shade of red that pushes through the skin barrier), but I got consumed with gratitude. I comprehended the fortune of being alive, healthy, within the comfort of my home, and, even lacking inspiration, having a purpose. A purpose that I simply knew I would be able to accomplish. Then I felt hope. Then, I could envision myself at this very present moment, writing my acknowledgements and thinking of the instants and steps I had built through my life in order to be here. I thought of every person who directly and indirectly offered the tools and grounds for me to dare to initiate and conclude my journey as a PhD student. In this part of my thesis, I would like to humbly thank all of those who honoured me with their generosity and light in various moments of my path within and beyond academia.

Above all, mama, thank you for welcoming me to this world and for your truly unconditional support and love in every single step I have ever taken. Your strength and kindness have forged all the good in me. There are no words capable of expressing how much you inspire me to be better. Thank you for existing and being the light of my life. I love you!

Foquinha, thank you for being, not only my partner in crime, but also the one who trained me for the most fun childhood mischiefs we could think of. Thank you for being a second mother when I needed, and my eternal best friend. Thank you for bringing Maria Helena into the world and into our lives, giving us hope and making us want to fight for a better future. I love until the end.

Papai, thank you for trying to be present despite the distance. I look up to your diligence and hard work. The example you set with your life have always incentivised me to give my very best in everything I do.

To my beloved grandparents and great-grandparents who I had the privilege to meet, thank you for being an example of humility and perseverance. Through your lives I could learn how to flexible and patient. To my aunts, uncles, half-siblings and cousins thank you for the joyful moments and tales we shared, I love each one of you tremendously. To the family of my partner who took me in and so actively supported me during the past years.

To all of my teachers, my dear ”mestres”… words truly fail me. Without you, I would be lost. I would not be. Thank you for teaching me the letters, the numbers, the colours and in a later moment for proving me with all necessary grounds to venture myself in any intellectual field. I would like to thank my professors at the Universidade Estadual da Paraíba, for teaching me how to be an ethical psychologist by trusting me with high stake tasks, and for allowing me to make mistakes so I could truly learn. In special, Professor Jorge Dellane, thank you for prioritising us students over the academic race, for reading our long essays, one by one, word by word carefully, and for writing us another essay back with your impressions about our work. You are one of a kind. Your generosity and love for teaching left a permanent mark in me. I will forever cherish your classes.

To my supervisor, Professor Henry Otgaar, thank you for believing in my potential to take over this great challenge, for your constant smile and reassurance. To my supervisor Dr. Sara Landtröm, thank you for your warm hugs and for always making me feel welcome; thank you for setting for me a powerful example of what it is to be a woman in academia, a wonderful mother, without ever forgetting to be true to yourself. Finally, my supervisor Dr. Timothy J. Luke, the rabbit I followed out of many holes. I was already some weeks through the programme when I got the news that a new supervisor would be joining my team. I still remember our first (skype) meeting and how you organically fitted in the group as if you had been there from the outset. As I got to know you, I knew that you would bring nothing but an excellent level of quality to my work. I wonder what my PhD would have turned out to be without you. As we moved forward with my projects, I also learned how your genuine and most unique generosity would impact me. Thank you for making me realise that it is okay to occupy space in people’s lives, for your infinite patience, and for presenting yourself to me in such an approachable manner. Thank you for sharing your passion for education, and for your unbreakable ethics. It is difficult to express how much I admire you and how grateful I am that you got rejected in all those hundred something jobs you applied for before you were invited to work in the University of Gothenburg.

To my dear friend Carey, thank you for daring to live with me and for being my family away from home. To Carey and Jenny thank you for the exhilarating adventures in Maastricht and abroad. To all the friends that this PhD quest brought to my life, including my programme peers and colleagues from the PhD academy, thank you for all the shared ideas and drinks, without your support and friendship this whole thing would have been boring and bleak.

To my dear friends in Brazil and other corners of the world, thank you for the long phone calls, and for not giving up on our friendship despite the challenging distance. I miss you every single day.

And lastly, to my partner Martien, thank you for daring to build a home with me, for your constant support, for feeding me delicious food, and for making me feel safe. I hope our bond continues to get stronger, and that we continue to humble ourselves before one another, learning from each other as we grow together. I love you.

## Curriculum Vitae

Bruna Calado, born on the 6th of November of 1991 in São José do Egito, Brazil has received her BSc and Lic degree in Psychology by the Universidade Estadual da Paraíba in Brazil. Bruna Calado is a qualified clinician in Psychology, as well as a specialised psychologist in Educational and Developmental Psychology. In 2017 Bruna Initiated her PhD in Legal Psychology hosted by the EMJD House of Legal Psychology programme at the host institution Maastricht University and second institution University of Gothenburg. Currently Bruna works with the development of technologies to foment high education at Maastricht University.


Harry Potter