City Research Online

City, University of London Institutional Repository

Citation: 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(3), pp. 365-378. doi: 10.1080/13218719.2016.1230924

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/15966/

Link to published version: https://doi.org/10.1080/13218719.2016.1230924

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Running head: CASE STUDY FALSE MEMORY

A Case Concerning Children's False Memories of Abuse:

Recommendations Regarding Expert Witness Work

Henry Otgaar[1,2], Corine de Ruiter[1], Mark L. Howe[1,2], Lisanne Hoetmer[1] , & Patricia van Reekum[1]

[1]Maastricht University, The Netherlands

[2]City University London, UK

Correspondence to Henry Otgaar, Faculty of Psychology and Neuroscience, Maastricht University, PO Box 616, 6200 MD, Maastricht, the Netherlands, Tel.: +31-43-3884340, Fax: +3143-3884196. E-mail address: Henry.Otgaar@maastrichtuniversity.nl

IN PRESS: *Psychiatry, Psychology, and Law*

Abstract

Expert witnesses can play a major role in legal cases concerning the reliability of statements. Abuse cases frequently contain only memories of eyewitnesses/victims without the presence of physical evidence. Here, it is of the utmost importance that expert witnesses use scientific evidence for their expert opinion. In this case report, we describe a case in which 20 children reported being sexually abused by the same teachers at their elementary school. We review the investigative steps that were taken by the police and school authorities and how they probably affected memory. In order to provide a sound expert opinion regarding the reliability of these statements, we propose three recommendations. So, to reduce the effect of confirmation bias and to increase objectivity, we argue that expert witness' reports should contain alternative scenarios, need to be checked by another expert, and should focus on the origin and context of the first statement.

Keywords: Legal case, Reliability, Eyewitness memory, False memory

A Case Concerning Children's False Memories of Abuse: Recommendations Regarding Expert Witness Work

Legal proceedings in which children testify that they have been sexually victimized are unique cases. These are frequently cases in which physical evidence (e.g., DNA) is absent and children's statements are diametrically opposed to the statements of the suspect (Brackmann, Otgaar, Sauerland, & Jelicic, 2016; Howe, 2013). In such proceedings, it is of vital importance to assess the reliability of children's statements. This is because the lynchpin in these cases is whether children's statements reflect an authentic experience or whether such accounts are tainted by suggestive influences and memory errors (i.e., false memories).

Among such child sexual abuse proceedings, there is a select group of cases in which a large number of children report having been maltreated by the same person(s) of an elementary school or daycare center (Garven, Wood, Malpass, & Shaw, 1998). During the 1980s, an "outbreak" of such "daycare abuse cases" occurred across the United States and Europe (Bottoms & Davis, 1997). The bottom line in many of these cases was that children often made false accusations based on false memories that arose during suggestive interviewing (for a recent review, see Howe & Knott, 2015). Oftentimes, such cases were characterized by bizarre sexual abuse allegations supposedly committed by a single perpetrator. In the last few years, this issue of these special abuse cases has gained renewed attention. Specifically, in many Western countries, similar cases have surfaced in which public figures were publically accused of having abused many children in previous decades (e.g., the Dutch Roman Catholic Church case; Saunders & Fivush, 2015). The central question in these cases is to estimate the reliability of statements of alleged (child) victims.

In the present case report, we describe a recent situation which has similar contours to these earlier child abuse proceedings. The research literature on children's memories for abuse has greatly expanded since the 1980s, so in theory the current expert witness should be able to provide a report with a much better empirical foundation compared to the experts of the 1980s. But applying research evidence to an individual legal case is less straightforward than it seems. The foremost aim of this case report is to arrive at some specific recommendations on how to deal with children's statements of abuse when (legal) psychologists are asked to act as expert witnesses. We start by explaining how children's statements can be adversely affected by the formation of false memories. Next, we describe some of the historic child abuse cases that involved false and suggested memories. We then describe the current case, explain how expert witness work is conducted in different jurisdictions, and provide recommendations for how future proceedings, should they arise, be better handled.

Eliciting False Memories

A burgeoning literature exists on ways of how to elicit false memories in the psychological laboratory. Of importance, these false memory methods have been invented to mimic real-life circumstances including, for example, suggestive police interviewing tactics. One of the most often-used methods to induce false memories is the misinformation paradigm (Loftus, 2005). In this paradigm, participants receive some stimuli (e.g., a video of a robbery) or are involved in an interactive event (e.g., participating in a magic show). Following this, they receive suggestive information (i.e., misinformation: "Was the culprit carrying a gun?" when no gun was present, or "Was the magician wearing a red cape?" when there was no cape) about the witnessed event. In the final stage, participants' memory for the original event is tested. The

standard finding is that many participants incorporate the misinformation into their memory reports thereby producing a false memory.

Although the misinformation paradigm elicits false memories for details, the implantation paradigm is known for its elicitation of rich false memories for entire events (e.g., Loftus & Pickrell, 1995; Otgaar, Candel, Merckelbach, & Wade, 2009). In this paradigm, participants are asked what they can remember about a fictitious event (e.g., a hot air balloon ride). Then, during multiple interviews, they have to report everything they can recollect about the false event. Research has shown that such suggestive interviews lead to 30-40% of participants falsely remembering that they experienced the false event (e.g., Otgaar, Scoboria, & Smeets, 2013; Wade, Garry, Read, & Lindsay, 2002).

What these false memory paradigms have in common is that they use suggestions about things that did not happen to evoke false memories. Recently, much attention has been directed towards understanding the production of *spontaneous* false memories that arise without external suggestive pressure (Brainerd, Reyna, & Ceci, 2008). A popular method used to induce spontaneous false memories is the Deese/Roediger-McDermott (DRM) paradigm (Deese; 1959; Roediger & McDermott, 1995). In the DRM paradigm, word lists containing associatively-related concepts (e.g., *fear, emotion, hate, fight*) are provided to participants. These concepts are connected to a common theme word called the critical lure (i.e., *anger*). A plethora of research has shown that many participants falsely remember the critical lure during subsequent free recall and recognition tests (Gallo, 2010).

The above-sketched false memory procedures have also been used with children. These studies have revealed that in many cases, children are exceptionally vulnerable to suggestive pressure and hence, the production of false memories (Bruck & Ceci, 1999; Otgaar, Candel,

Smeets, & Merckelbach, 2010). Furthermore, many of these studies have shown that false memories induced by suggestion are more likely to occur in younger children than in older children and adults (Ceci & Bruck, 1993; Otgaar & Candel, 2011; but see Otgaar, Howe, Brackmann, & Smeets, 2016). In contrast to this age-related pattern for suggestibility, developmental studies on spontaneous false memories have demonstrated that adults are more likely to create false memories than children, a phenomenon dubbed "developmental reversal" (Brainerd et al., 2008).

In legal cases, both suggestion-induced and spontaneous false memories can surface in both children's and adults' statements. Both children and adults can be suggestively interviewed by parents, therapists, or police officers and hence, such interviews might negatively affect their statement quality. Although research shows that when children (or adults) spontaneously report having been maltreated, their statements often possess a high level of accuracy (e.g., Bidrose & Goodman, 2000; Leander, Christianson, & Granhag, 2007; Orbach & Lamb, 1999), recent work on false memories shows that adults (and to a lesser extent children) can be prone to producing spontaneous memory errors in their reports. Understanding how such false memories arise is crucial for legal professionals as they often have incorrect ideas about how memory works (Odinot, Boon, & Wolters, 2015; Ost, Easton, Hope, French, & Wright, in press). A specific example of this is that legal professionals sometimes incorrectly assume that children are inferior eyewitnesses because of their extreme proneness to false memories, thereby discrediting statements of young children (for a recent case report on this issue, see Brackmann et al., in press).

False Memory Cases

Although during the 1980s the number of cases of alleged abuse at daycares increased considerably, only a few of them sparked extensive attention among memory researchers, legal professionals, and the media. One prominent case was the McMartin Preschool trial (Garven et al., 1998; Garven, Wood, & Malpass, 2000). In 1983, seven teachers from the McMartin Preschool located in Los Angeles were charged with sexually abusing more than 100 children in the most horrifying manner. Children "remembered" some extremely bizarre experiences such as seeing dead and burned babies, riding naked on a horse, and flying to an isolated farm where they had to engage in group sex (Schreiber et al., 2006).

The accuracy of the children's reports was highly contested because of several contaminating influences. First, after the first allegation, parents received a letter in which they were asked to question their children to see whether they were victimized. There is research suggesting that parents often ask children suggestive questions when abuse is suspected (Korkman, Juusola, & Santtila, 2014). Second, children were also interviewed in a highly suggestive manner by social service workers, which might have led children to make false statements. An abundance of research reveals that children are particularly susceptible to suggestive questioning and hence, the formation of false memories (Bruck & Ceci, 1999). Most of the charges against the teachers were dropped eventually and none of the teachers was convicted of a sexual crime.

Another highly publicized daycare abuse case was one that occurred at the Wee Care Nursery (Ceci & Bruck, 1993). In this case, nursery school teacher Margaret Kelly Michaels was accused of sexually abusing children at the Wee Care Nursery School. Like the McMartin Preschool case, some of the allegations made by the children were quite bizarre. For example, she was accused of: Licking peanut butter off children's genitals, playing the piano while being

nude, making children drink her urine and eat her feces, and raping children with Lego blocks, knives, forks, and spoons. According to the children's accounts, she abused them during school hours over a period of seven months.

As in the McMartin Preschool case, there are several elements in this case that likely infected children's accounts of sexual abuse. First, children's parents received a letter from the daycare center informing them that some children had stated they were abused. Second, during a subsequent meeting, a social worker told parents that sexual abuse of children is very common and that the parents should examine their children for signs of sexual abuse (e.g., nightmares, bed wetting). After this, the number of allegations against Kelly Michaels increased significantly. Kelly Michaels was sentenced to 47 years imprisonment but was released after 5 years because a number of legal psychologists sent an amicus letter to the court expressing concerns regarding the accuracy of the children's sexual abuse statements (Bruck & Ceci, 1995).

Daycare abuse cases are not confined to the United States. Similar cases have appeared in European countries as well. For example, in 1987, in the small Dutch village of Oude Pekela, children reported to have been sexually molested by a group of satanic clowns (Goodman et al., 1997; Jonker & Jonker-Bakker, 1991). In this case, the idea was that these accusations were the result of suggestive interviewing tactics. Other notable examples are the Wormser and Montessori trials in Germany during the 1990s in which children's reports of abuse were likely influenced by suggestive interviewing techniques as well (Schade & Harschneck, 2000). In the case that we describe here, and in which two of the authors were involved as expert witnesses (first (HO) and second (CdR) authors), we analyze a recent child sexual abuse case that has many similarities with previous daycare abuse cases.

The "Galileo" Case

In this case[1], HO and CdR were contacted by parents of alleged child victims of sexual abuse. The parents notified them that their children had been sexually abused at an elementary school called "Galileo." This school was located in a small town in the western part of the Netherlands. In September 2009, a father of one of the alleged victims filed an official charge with the police on behalf of his eight-year old son named Guy. The father declared that his son had been sexually abused by two teachers of the "Galileo" elementary school. After the charge, the police interviewed more children and they reported having been abused as well. In total, 20 children reported having been abused by the same two teachers of the school.

When reading the case file, a number of critical facts became evident. In cases such as this, it is essential that experts examine how the first statement of abuse was produced (see e.g., Volbert & Steller, 2014). This information is important because for children, statements that occur spontaneously are often provided without any (suggestive) intervention and may provide a more accurate account of what was experienced (e.g., Bidrose & Goodman, 2000). Furthermore, it is not uncommon that such spontaneous statements are provided after long delays as child victims of sexual abuse are often reluctant to talk about their abusive experiences (London, Bruck, Ceci, & Shuman, 2005). Interestingly, in the Galileo case, the first statement that was provided by the child had nothing to do with sexual abuse. That is, in 2008, Guy stated that a female teacher named Maria had punished him in class. Guy talked about this to his parents. Specifically, Guy was at times not allowed to go to lunch or to the toilet. Maria confirmed that this had happened and other teachers also reported that Maria was a strict teacher.

Because Guy had many problems in and out of school (i.e., he was highly anxious about going to school), he had been undergoing therapy. The therapist concluded that Guy had

[1] Names in this case have been changed to protect the involved parties

experienced something highly traumatic and that he had Posttraumatic Stress Disorder. To uncover the traumatic experiences, the therapist used anatomically-detailed dolls and Guy was asked to show whether he was perhaps (sexually) abused. A large body of research shows that such dolls are highly suggestive, lead to fantasy play behavior, and increase the chance of incorrect reports (e.g., Otgaar, Horselenberg, Van Kampen, & Lalleman, 2012; Poole & Bruck, 2012). Hence, to date, the use of such props is discouraged in interviews with child victims of alleged sexual abuse (Bruck, 2009; Poole, Bruck, & Pipe, 2011; see also Otgaar, Van Ansem, Pauw, & Horselenberg, in press). Although Guy did not come up with any details concerning sexual abuse during the therapy session with the anatomically-detailed dolls, it was only after this (suggestive) therapeutic session that Guy started to change his statement about punishment into statements concerning sexual abuse.

One day after the first official police charge, the school sent a letter to all parents about possible abuse at the elementary school. In this letter (see Appendix for the text of the letter), the parents were informed of this alleged abuse and told that more children could have been victimized. Moreover, the parents were told that they would receive information on how to interpret signs that their child was abused. After the letter, many other children came forward, having "remembered" being sexually abused at the school. It is striking that in many earlier daycare abuse cases (e.g., McMartin Preschool, Wee Care Nursery School), such (suggestive) letters had also been sent to parents prior to their children making similar accusations of having been sexually abused. Also, it is important to remember that the idea that there are specific signs indicative of (sexual) abuse has not received any empirical support. In fact, research shows that there are no reliable signs that are diagnostic of child sexual abuse (Kelly, Koh, & Thompson, 2006; Kendall-Tackett, Williams, & Finkelhor, 1993).

In the present case as in the earlier ones, the children's statements contained certain bizarre details. For example, Guy reported that Maria had cut off a piece of his penis. Furthermore, some children remembered that they had to go to one of the teachers' houses where they had to look at crocodiles. Among some clinicians, there is the idea that children do not make up such bizarre stories and that hence, when bizarre details do arise, they must be authentic (Dahlenberg, 1996; Everson, 1997). However, research shows that false memories for highly implausible events are can be easily evoked even in children (Otgaar et al., 2009; see also Mazzoni, Loftus, & Kirsch, 2001).

A final remark pertains to the fact that many of the children's parents talked on many occasions about the alleged sexual abuse. They even held a meeting to discuss their concerns. Numerous studies have shown that such discussions might inadvertently affect children's statements (e.g., Principe & Schindewolf, 2012). More precisely, these studies indicate that discussions might lead to rumours concerning, for example, sexual abuse. When children are subjected to such rumours, this can lead them to falsely remember that they too were abused (Principe & Schwindewolf, 2012).

Taken together, this case has many parallels with previous daycare abuse cases. Although Guy did spontaneously report experiences concerning (harsh) punishment, his subsequent statements are likely to have been adversely affected by what we consider to be obvious indications of suggestive influence. These include the distribution of the letter to all parents by the school alleging sexual misconduct and the suggestive methods used by Guy's therapist. At first, the prosecution did not prosecute the case because of a lack of convincing proof that these crimes had taken place and because a legal psychologist wrote a report casting doubt on the reliability of the children's statements. Following this, the parents contacted HO and CdR and

went to the Court of Appeal. The Court of Appeal decided that because the expert witnesses concluded that the children's statements could have been affected by suggestive influences, the teachers would not be prosecuted. In cases like the one described here, expert witnesses have a crucial and often decisive role in determining whether the statements are reliable. Hence, expert witnesses should look carefully at all case files in order to provide an objective judgment about the reliability of children's statements.

Expert Witness Work in Different Countries

The "Galileo" case took place in the Netherlands. This is important to point out because expert witnesses can play somewhat different roles depending on which country they are in and which criminal legal system pertains to that country. Broadly speaking, criminal law systems from across the world can be divided into two types: Inquisitorial and adversarial (De Ruiter & Kaser-Boyd, 2015). The Netherlands uses the strictest inquisitorial system in the world whereas countries like England and the United States adopt an adversarial approach. In the adversarial system, legal trials can be regarded as contests between equivalent parties (Van Koppen, 2007): the prosecution and defense. They argue their case in a somewhat equal role in front of the judge (and oftentimes, the jury). Under the adversarial model, the judge is considered to have a neutral position, intervenes only in the case of conflicts, and is a purveyor of all things to do with the law.

The adversarial system is different from the inquisitorial one. Here, there is no jury. Also in, for example the Netherlands, a trial is not about a fair contest between parties, but about finding the truth (Van Koppen & Penrod, 2003). The judge is a professional judge in that compared to the adversarial system, judges in for example the Netherlands receive special training. They have to be in law school for five years and then undergo an additional six years of

special judicial training. Part of this extra training involves courses concerning the psychology of law (Van Koppen, 2007). Furthermore, in the inquisitorial system, the judge has an active role and leads the investigation. Furthermore, the prosecution aims to come up with an independent judgment before a case can be continued and this can mean that a prosecutor asks for an acquittal. Also under this model, defense attorneys and prosecutors are not the ones primarily asking questions in the courtroom. Because the judge has an active role, he/she will be primarily in charge of asking questions of the defendant.

Expert witnesses have a rather unique position in each of these legal systems. That is, in both systems, expert witnesses can be called upon to appear in court and in both systems, expert witnesses can be approached by either the prosecution or the defense. However, one crucial difference is that in the adversarial system, there is the chance that expert witnesses are somewhat biased to the side that retained them (Saks, 2003). Indeed, recent research has shown that such allegiance biases occur in expert witnesses (Murrie, Boccaccini, Guarnera, & Rufino, 2013). Such biases might lead to battles between experts that were hired by opposing parties. Although in the inquisitorial system, expert witnesses are assumed to be more independent (Saks, 2003) and are frequently not asked to elaborate on their expert witness report in the courtroom, this does not prevent them from battling with other experts in the courtroom.

To illustrate, we have recently described a case in which the first author (HO) was asked to evaluate the reliability of a series of statements made by a 6-year-old girl (Brackmann et al., in press). The basic conclusion of this expert report was that there were no strong reasons to doubt the statements of the girl. The defense had hired another expert, a clinical psychologist, who reasoned that the girl's statements were contaminated because of autosuggestion (i.e., similar to what we discussed earlier in terms of spontaneous false memories that arise without any external

suggestive pressure). However, during the trial, the point was made by HO that recent research shows that autosuggestive effects are less likely to occur in young children than in older children and adults (Brainerd et al., 2008) and that hence, autosuggestion was unlikely to play a role in the statements of this child. Eventually, this resulted in a disagreement between the two experts in the courtroom concerning the susceptibility of false memory production in children and adults.

Of course, part of the reason why there may be disagreement between experts is related to the fact that experts are often experts in different domains (e.g., the science of memory vs. clinical psychology) and issues can arise as to which expert testimony is admitted in court. For example, in the United States, only expert evidence that is allowed is that which meets the Frye or Daubert criteria (e.g., known error rate, subjected to peer review and publication; Saks & Faigman, 2005). Although such specific criteria are absent in the Netherlands, expert databases are currently being constructed, ones that contain expertise in various fields (e.g., forensic psychology) and that have been demonstrated to be accepted by the scientific community. Nonetheless, in both systems, it remains possible that expert witnesses are hired who lack the scientific knowledge about the subject they are asked comment on.

Before we were asked to review the Galileo case, there was another report by a legal psychologist. The parents argued that the first psychologist had not looked at all of the necessary case files in order to provide an appropriate and objective conclusion. Hence, after the parents contacted us, we decided not to look at the earlier report to prevent ourselves from developing any expectancy effects about the case and looked at all of the case files independently. Finally, when we wrote our expert witness report and then compared it with the previous one, we found many similar arguments and a highly similar conclusion that there were elements in the case that could have affected the children's reports.

Based on this case and other expert witness work in which we were involved, we believe it is important to strengthen expert witness decisions by offering a number of recommendations. It is our contention that the following recommendations need to be implemented by expert witness in general, whether they are working in adversarial or inquisitorial systems.

Recommendations

In the majority of child sexual abuse cases, the central question revolves around whether statements are reliable or not. In order to provide a well-balanced conclusion concerning the reliability of statements, we propose a set of three recommendations. Our stance is that when adopting these recommendations, expert witnesses are in a better position to provide an unbiased conclusion based on the latest scientific findings.

Our first recommendation is that an expert witness should commence with so-called "scenario building" (Crombag & Wagenaar, 2000; Raskin & Esplin, 1991; Rassin, 2014). Our proposal is that when starting with expert witness work concerning the reliability of testimony, one should postulate at least two scenarios. One scenario would signify the "guilty" scenario in which the assumption is that statements refer to an authentic experience.[2] The alternative scenario could be called the "innocent" scenario in which the assumption is that statements are the result of fabrication (e.g., false memories, deception). Obviously, such a recommendation is not only relevant for expert witnesses. More generally, it should be used throughout the entire investigative process and is recommended for the police during the interview process as well.

The idea behind this is that when expert witnesses are working on a case concerning the reliability of statements, they should decide whether they find elements in the case supporting the "guilty" or "innocent" scenario. If they find more support for the "innocent" scenario, then

[2] The terms "guilty" or "innocent" scenario do not imply that the expert witness has to decide whether the suspect is guilty or not. These terms for the different scenarios have been used frequently by Dutch legal psychologists (e.g., Van Koppen, 2011) and refer to whether statements are based on true experiences or are fabrications.

there are reasons to doubt the reliability of the statements. However, if more support is found for the "guilty" scenario, then there are few reasons to doubt the reliability of statements. To give an example, when reading a case file concerning the reliability of a suspect's confession, the expert witness discovers that the suspect only confessed to the crime after a suggestive interrogation by the police. Because research shows that this might lead to false confessions (Kassin, 2008), the expert witness has found a piece of evidence in support for the alternative "innocent" scenario.

Why the construction of such scenarios is important is because it might protect the expert witness from expectancy effects or confirmation bias (Kassin, Dror, & Kukucka, 2013). Specifically, by using such scenarios, the expert witness is not solely focused on, for example, finding evidence for proof of false memories or the reverse. Also, it might guard expert witnesses from something called the allegiance bias in which forensic experts are unconsciously biased towards the party that they are retained by (i.e., prosecutor vs. defense; Murrie et al., 2013). Indeed, although not in the realm of expert witness work, there is evidence suggesting that the postulation of alternative hypotheses immunizes against confirmation bias. Specifically, O'Brien (2009; Study 2) provided participants with a police file of a criminal investigation and their task was to answer certain questions about the case. Participants were assigned to different conditions, but of relevance here is the condition in which participants were asked who committed the crime (generating a hypothesis) and the condition in which participants also had to think about why the person might be innocent (alternative hypothesis). O'Brien found that statistically significant more bias was observed in the generate-a-hypothesis than in the alternative-hypothesis group inasmuch as participants in the generate-a-hypothesis group remembered more facts consistent with the guilt of the suspect than the alternative-hypothesis group.

Recent research has also shown that thinking about alternatives might make jurors more critical for eyewitness evidence. That is, Rodriguez and Berry (in press) provided mock jurors with a brief crime summary containing high- (e.g., unbiased lineup instructions) or low-quality (e.g., biased lineup instructions) eyewitness evidence. Then, half of the participants had to adopt a counterfactual thinking style (i.e., considering how the present situation could have been different) which is closely linked to coming up with an alternative scenario. The other half did not have to engage in counterfactual thinking. The most important result was that when participants received low-quality eyewitness evidence, there were least likely to reach guilty verdicts when they used a counterfactual mindset. Collectively, these studies suggest that alternative scenarios might lead people to search for different outcomes thereby potentially reducing confirmation bias.

There have been attempts to increase the relevance of scenario building (Volbert & Steller, 2014). Although these attempts are important, they have been rather unsuccessful. This is because expert witnesses are often unclear exactly when this approach should be used and how many scenarios should be constructed (Neal & Brodsky, 2016). In our view, scenario building should be done *before* reading the case files. So, in its simplest form, only two scenarios should be proposed at the start of the case – one referring to statements based on a true experience and one indicative of a false statement. Of course, the term "false" can suggest different issues such as intentional fabrication (i.e., lies) or unintentional fabrication (i.e., false memories). However, this nuance is not relevant at the beginning, as expert witnesses commonly have no prior knowledge concerning the case. After reading the entire case file, additional (exploratory) scenarios can be generated that provide more nuances in scenario building. To give a concrete example, suppose expert witness X is asked to provide his/her expert opinion on statements of

several eyewitnesses of a crime (e.g., theft). Expert X postulates two scenarios (guilty vs. innocent). When reading the case files, expert X discovers that the supposed eyewitnesses have had multiple arguments with the suspect before the theft and that they were not actually present during the crime. Also, expert X finds out that the eyewitnesses have had many meetings with each other to discuss the case. These particular elements in this case might lead to an additional (exploratory) scenario referring to idea that the eyewitness statements are not only false but have been affected by deception as well.

The idea of postulating scenarios before reading the case files comes close to the concept of preregistration in science (Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). Because of the increased attention on failed replications in psychology, the idea of preregistration is gaining momentum in the psychological field. The idea behind preregistration is quite easy; researchers indicate before conducting an experiment the hypotheses and planned analyses on a public forum. After data collection, exploratory analyses can be conducted but the preregistered hypotheses (and analyses) are the most relevant for the experiment and fall under the label, confirmatory research. To some extent, this is exactly what is done during scenario building in expert witness work. Admittedly, it is true that no empirical knowledge yet exists on whether the use of this preregistration strategy results in better expert witness work. However, such hypotheses would be good to test in future research.

The model of "scenario building" is also highly linked with hypothesis testing in many scientific disciplines. That is, before conducting an experiment, it is advised to formulate a null hypothesis and an alternative hypothesis (Gigerenzer, 2004; Schneider, 2015). The null hypothesis is analogous to our "guilty" scenario and the alternative hypothesis refers to the "innocent" scenario. Here too, when enough evidence exists for the alternative hypothesis (e.g., *p*

< .05), the null hypothesis is rejected. Also, scenario building bears many similarities to Bayesian hypothesis testing. In Bayesian hypothesis testing, one can calculate average likelihoods also known as Bayes Factors. Bayes Factors refer to the degree to which the data are more likely under the null than under the alternative hypothesis (Wagenmakers et al., in press). To some extent, this is exactly what scenario building entails. When using scenario building, the task of the expert witness is to judge whether a given case file is more likely to contain elements for the "guilty" or "innocent" scenario.

In the case outlined here, it is evident that there is support for the "guilty" scenario in that Guy spontaneously talked about being harshly punished and this was confirmed by his teacher and her colleagues. However, the case also contained a number of elements (e.g., (suggestive) letter, meeting with parents, suggestive therapy) that spoke in favor of the "innocent" scenario. Taken together, our conclusion was that besides the spontaneous statement of Guy regarding harsh punishment, there were reasons to doubt the reliability of the statements concerning sexual abuse.

Our second recommendation is that after setting up different scenarios, the expert witness should focus on the circumstances around the *first* statement. This is critical information as studies have revealed that spontaneous statements often contain a high degree of accuracy (Orbach & Lamb, 1999). Furthermore, when there is information in the case file that the first statement arose after suggestive interviewing, this is also essential information to estimate the reliability of the statement. This is especially true when a case consists of repeated interviews. Repeated interviews by themselves are not necessarily dangerous for the reliability of eyewitness memory (Goodman & Quas, 2008). However, when the first statement is developed after suggestive pressure, then follow-up statements are likely to contain erroneous details even if

these statements were the result of well-conducted interviews (Ceci, Kulkofsky, Klemfuss, Sweeney, & Bruck, 2007).

A third recommendation for expert witness work is that the draft expert witness report should be read by another blind and unbiased expert. Ideally, such an expert witness should be someone who is also an expert in the field of, for example, eyewitness memory and who has no knowledge concerning the specifics of the case. The task of this person is to review the report and to provide critical comments where appropriate. This is an additional strategy to minimize the contaminating influence of confirmation bias. Therefore, it is important to note that this second expert is not merely a proof-reader. On the contrary, the task of this second expert is to critically check whether the expert witness has carefully used the scientific literature and the information in the case file to support his/her scenarios and conclusion. One might think that having an extra expert review the report amounts to too much work and is therefore not feasible. However, this approach is common practice among Dutch legal psychologists. Another reason why the expert report should be read by a second expert is because the act of scenario building can lead to some subjective judgments by the expert witness. For example, it is not uncommon that an expert witness finds some support for the "innocent" (e.g., suggestive interviewing) *and* "guilty" (e.g., first statement was spontaneous) scenario and then the question is how the expert witness should weigh these findings. A second expert can test the manner by which such judgments are made and examine whether the decision concerning the weighing has any face value.

It is our opinion that when using these three rather straightforward recommendations, expert witnesses guard themselves against making faulty inferences and unintentionally focusing more on one specific scenario. The downside to conclusions based on scenario building is that

they are mostly qualitative (i.e., “there are no reasons to doubt the reliability of statements”). This results in conclusions being somewhat superficial as there are no clear cut-off scores for what constitutes an unreliable statement. Of course, there are some exceptions to this disadvantage as, for instance, levels of suggestibility can be quantitatively assessed using the Gudjonsson Suggestibility Scale (GSS; Gudjonsson, 1984) and high levels of suggestibility might be seen as somewhat supportive for the “innocent” scenario. Nonetheless, the method of scenario building – which is widely used among Dutch legal psychologists (e.g., Rassin, 2014) – could benefit from a more quantitative approach. Rassin suggested scoring each element (e.g., suggestive interviews) in favor of the “innocent” scenario with 1 point. The more points a case contains, the more likely the statements are contaminated. Although others have suggested similar strategies as well (O’Donehue, Benuto, & Cirlugea, 2014), the problem with such an approach is how much weight should be given to each element. For example, if an eyewitness reports having seen the perpetrator of a robbery at night, it is difficult to decide how much weight should be given to the bad lighting conditions, especially if there are no other reasons to doubt the statements. Or consider the following: Two adult sisters claim to have been sexually abused by their father when they were six years old. One of the sisters came forward with the statement one year after the alleged event while the other sister took eight years to come forward. For both sisters, their first statement occurred after a delay, but it is quite difficult to decide when a delay can be seen as a reason to doubt the statement or can be regarded as typical for child sexual abuse cases (i.e., delayed disclosure; e.g., Goodman-Brown, Edelstein, Goodman, Jones, & Gordon, 2003). Future research is needed to examine whether such a quantitative approach has scientific merit.

Obviously, the recommendations that we have just put forward are specifically focused on expert witness work. However, other parties could benefit from such recommendations as well. The model of scenario building is relevant for everyone interviewing alleged victims of sexual abuse (e.g., police, therapists). That is, when taking into consideration that statements might also be false, statements are not routinely accepted as true and will be examined more critically. This might reduce the likelihood that suggestive influences occur in child sexual abuse cases. For example, it might reduce the tendency to distribute (suggestive) letters to parents by schools. In another recent abuse case in which one of us worked as an expert witness (HO), the director of the elementary school knew about the dangers of suggestive influences on memory. In that case, the school also sent a letter but no mention was made about the possibility of other children being victimized. Also, correct advice was provided on how children should be interviewed.

To recap, we have described a recent elementary school abuse case, one that has many similarities to the classic daycare abuse cases from the 1980s and 1990s. From this, we have argued that expert witnesses play a pivotal role in advising courts (whether the system is inquisitorial or adversarial) and that to increase their objectivity and transparency, some simple and easily applicable recommendations should be used during the course of their preparation of an expert report. By embracing these recommendations, not only will such reports be more objective and standardized, but this approach will also enhance the scientific status of psychologists in the courtroom.

References

Bidrose, S., & Goodman, G.S. (2000). Testimony and evidence: A scientific case study of memory for child sexual abuse. *Applied Cognitive Psychology, 14*, 197-213.

Bottoms, B.L., & Davis, S.L. (1997). The creation of satanic ritual abuse. *Skeptical Intelligences, 3*, 1-12.

Brackmann, N., Otgaar, H., Sauerland, M., & Jelicic, M. (2016). When children are the least vulnerable to false memories: A true report or a case of autosuggestion? *Journal of Forensic Sciences, 61,* S271-S275.

Brainerd, C.J., Reyna, V.F., & Ceci, S.J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134,* 343-382.

Bruck, M. (2009). Human figure drawings and children's recall of touching. *Journal of Experimental Psychology: Applied, 15*, 361-374.

Bruck, M., Ceci, S.J. (1995). Amicus brief for the case of *State of New Jersey v. Michaels* presented by committee of concerned social scientists. *Psychology, Public Policy and Law, 1*, 272-322.

Bruck, M., & Ceci, S.J. (1999). The suggestibility of children's memory. *Annual Review of Psychology, 50*, 419-439.

Ceci, S.J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113,* 403-439.

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.

Crombag, H.F.M., & Wagenaar, W.A. (2000). Audite et alteram partem. *Trema,* 93-96.

Dahlenberg, C.J. (1996). Fantastic elements in child disclosures of abuse. *APSAC Advisor, 9,* 1-10.

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

De Ruiter, C., & Kaser-Boyd, N. (2015). *Forensic psychological assessment in practice: Case studies*. Routledge: New York and London.

Everson, M. (1997). Understanding bizarre, improbable, and fantastic elements in children's accounts of abuse. *Child Maltreatment, 2,* 134-149.

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

Garven, S., Wood, J.M., & Malpass, R.S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology, 85,* 38-49.

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

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics, 33,* 587-606.

Goodman, G.S., & Quas, J.A. (2008). Repeated interviews and children's memory. It's more than just how many. *Current Directions in Psychological Science, 17,* 386-390.

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.

Goodman-Brown, T.B., Edelstein, R.S., Goodman, G.S., Jones, D.P.H., & Gordon, D.S. (2003). Why children tell: A model of children's disclosure of sexual abuse. *Child Abuse & Neglect, 27*, 525-540.

Gudjonsson, G.H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences, 5*, 303-314.

Howe, M.L. (2013). Memory lessons from the courtroom: Reflections on being a memory expert on the witness stand. *Memory, 21,* 576-583.

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

Jonker, F., & Jonker-Bakker, P. (1991). Experiences with ritualist child sexual abuse: A case study from the Netherlands. *Child Abuse & Neglect, 15,* 191-196.

Kassin, S.M. (2008). False confessions. Causes, confessions, and implications for reform. *Current Directions in Psychological Science, 17*, 249-253.

Kassin, S.M., Dror, I.E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition, 2*, 42-52.

Kelly, P., Koh, J., & Thompson, J.M. (2006). Diagnostic findings in alleged sexual abuse: Symptoms have no predictive value. *Journal of Paediatrics and Child Health, 42*, 112-117.

Kendall-Tackett, K.A., Williams, L.M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113*, 164-180.

Korkman, J., Juulosa, A., & Santtila, P. (2014). Who made the disclosure? Recorded discussions between children and caretakers suspecting child abuse. *Psychology, Crime & Law, 20,* 994-1004.

Leander, L., Christianson, S.A., & Granhag, P.A. (2007). A sexual abuse case study: Children's memory and reports. *Psychiatry, Psychology and Law, 14,* 120-129.

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

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

London, K., Burck, 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.

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.

Murrie, D.C., Boccaccini, M.T., Guarnera L.A., & Rufino, K.A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science, 24,* 1889-1897.

Neal, T.M.S., & Brodsky, S.L. (2016). Forensic psychologists' perceptions of bias and potential correction strategies in forensic mental health evaluations. *Psychology, Public Policy, and Law, 22,* 58-76.

O'Brien, B. (2009). Prime suspect: An examination of factors that aggravate and counteract confirmation bias in criminal investigations. *Psychology, Public Policy, and Law, 15,* 315-334.

O'Donehue, W., Benuto, L.T., & Cirlugea, O. (2014). Analyzing child sexual abuse allegations: Further considerations. *Journal of Forensic Psychology Practice, 14*, 237-246.

Odinot, G., Boon, R., & Wolters, L. (2005). Het episodisch geheugen en getuigenverhoor. Wat weten politieverhoorders ervan? *[Episodic memory and eyewitness interviewing. What do police interrogators know about this?] Tijdschrift voor Criminologie, 57*, 279-299.

Orbach, Y., & Lamb, M.E. (1999). Assessing the accuracy of a child's account of sexual abuse: A case study. *Child Abuse & Neglect, 23*, 91-96.

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

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

Otgaar, H., Candel, I., & Merckelbach, H. (2008). Children's false memories: Easier to elicit for a negative than for a neutral event. *Acta Psychologica, 128*, 350-354.

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.

Otgaar, H., Candel, I., Smeets, T., & Merckelbach, H. (2010). "You didn't take Lucy's skirt off": The effect of misleading information on omissions and commissions in children's memory reports. *Legal and Criminological Psychology, 15*, 229–241.

Otgaar, H., Horselenberg, R., van Kamper, R., & Lalleman, K. (2012). Clothed and unclothed human figure drawing lead to more correct and incorrect reports of touch in children. *Psychology, Crime and Law, 18*, 641-645.

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.

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.

Otgaar, H., Van Ansem, R., Pauw, C., & Horselenberg, R. (in press). Improving children's interviewing methods? The effects of drawing and practice on children's memories for an event. *Journal of Police and Criminal Psychology.*

Poole, D.A., & Bruck, M. (2012). Divining testimony? The impact of interviewing props on children's reports of touching. *Developmental Review, 32,* 165-180.

Poole, D.A., Bruck, M., & Pipe, M. (2011). Forensic interviewing aids: Do props help children answer questions about touching? *Current Directions in Psychological Science, 20,* 11-15.

Principe, G.F., & Schindewolf, E. (2012). Natural conversations as a course of false memories in children: Implications for the testimony of young witnesses. *Developmental Review, 32,* 205-223.

Raskin, D.C., & Esplin, P.W. (1991). Statement validity assessment: Interview procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment, 13,* 265-291.

Rassin, E. (2014). Het beoordelen van de validiteit van (getuigen)verklaringen met de methode van de alternatieve scenario's; een update *[Judging the validity of (eyewitness) statements with the method of alternative scenarios; an update]*. *Expertise en Recht, 4,* 119-123.

Rodriguez, D.N., & Berry, M.A. (in press). Sensitizing potential jurors to variations in eyewitness evidence quality using counterfactual thinking. *Applied Cognitive Psychology.*

Roediger, H. L., & McDermott, K. (1995). Creating false memories: Remembering words not presented in a list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Saks, M.J. (2003). Expert witnesses in Europe and the United States. In P.J. Van Koppen and S.D. Penrod (Eds)., *Adversarial versus inquisitorial justice: Psychological perspectives on criminal justice systems*. Springer Science & Business Media.

Saks, M.J., & Faigman, D.L. (2005). Expert evidence after Daubert. *Annual Review of Law and Social Science, 1,* 105-130.

Saunders, J., & Fivush, R. (2015). Reshaping memories through conversations: Considering the influence of others on historical memories of abuse. *Applied Cognitive Psychology, 29,* 789-790.

Schade, B., & Harschneck, M. (2000). Die BGH-Entscheidung im Rückblick auf die Wormser Missbrauchsprozesse. Konsequenzen für die Glaubhaftigkeitsbegutachtung aus der Sicht des psychologischen Gutachters und des Strafverteidigers *[The Federal Court of Justice decision in retrospect of the Wormser abuse trials. Consequences for the credibility assessment from the perspective of the psychological expert and the defence lawyer*]. *Praxis der Rechtspsychologie, 10*, 28-47.

Schneider, J.W. (2015). Null hypothesis significance tests. A mix-up of two different theories: The basis for widespread confusion and numerous interpretations. *Scientometrics, 102,* 411-432,

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 daycare abuse cases: A case study. *Social Influence, 1,* 16-47.

Van Koppen, P.J. (2007). Miscarriages of justice in inquisitorial and accusatorial legal systems. *The Journal of the Institute of Justice & International Studies, 7,* 50-58.

Van Koppen, P.J., & Penrod, S.D. (2003). Adversarial or inquisitorial: Comparing systems. In P.J. Van Koppen and S.D. Penrod (Eds)., *Adversarial versus inquisitorial justice: Psychological perspectives on criminal justice systems*. Springer Science & Business Media.

Volbert, R., & Steller, M. (2014). Is this testimony truthful, fabricated, or based on false memory? Credibility assessment 25 years after Steller and Koehnken (1989). *European Psychologist, 19,* 207-220.

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

Wagenmakers, E.-J., Verhagen, A. J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (in press). The need for Bayesian hypothesis testing in psychological science. In Lilienfeld, S. O., & Waldman, I. (Eds.), *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*. Chichester, UK: John Wiley and Sons.

Wagenmakers, E., Wetzels, R., Borsboom, D., van der Maas, H.L.J., & Kievit, R.A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7,* 632-638.

Appendix

Translated and edited version of the school's letter sent to the parents.

Dear parents,

On behalf of the Galileo school, we want to invite you for a special meeting tomorrow evening. We want to inform you about an official charge that a parent made to the police against two staff members of our school. In this letter, we want to inform you about what happened, what the status is of the police investigation, and how we can help you.

What happened

Last Monday, we were confronted with a parent that stated that two staff members had inappropriate physical contact with their child. The parent also filed a charge at the police about this and the police has now started an investigation concerning the charge.

Information for parents

We are deeply shocked about this event. We have to wait for the outcome of the investigation but the accusations are of such a nature that we want to inform you. We want to do this because other children are possibly involved in this case as well.

Support

We are aware that this letter might lead to concern and questions regarding your child. During the meeting, we will inform you about the current status of this case. Professionals will explain how to interpret signs and signals of your child and how you can help your child. Furthermore, advice will be given about what not to do.