CONSIDERATIONS FOR EXPERTS IN ASSESSING THE CREDIBILITY OF RECOVERED MEMORIES OF CHILD SEXUAL ABUSE
The Importance of Maintaining a Case-Specific Focus

Laurence Alison  
University of Liverpool

Mark Kebbell  
Griffith University

Penney Lewis  
King’s College London

In this article, the authors argue that a variety of psychological factors stand in the way of providing expert advice to the courts in terms of assessing the credibility of a complainant’s account of sexual abuse when there is a significant delay in reporting. These include difficulties in assessing (a) the complainant’s account of how he or she claims to have remembered or forgotten the abuse, (b) whether (and how) the claim of abuse originated within a therapeutic setting, and (c) the difficulty of generalizing from empirical evidence. It is argued that all of these issues can be more easily avoided if experts maintain a case-specific focus. In this article, the authors review both the psychological and legal controversies surrounding the false–recovered memory debate, discuss how courts approach the admissibility and use of recovered memory testimony, and conclude that expert witnesses should carefully consider the above points before drawing general conclusions from the literature and applying them to individual cases.

Keywords: child sexual abuse, recovered memories, sexual abuse victims, expert testimony

Recovered memories have been classified as cases in which adults initially believe they were not sexually victimized as children and later come to believe that they were, rather than cases in which people who always knew they survived such abuse as children remember additional details or instances. (Lindsay & Read, 1995, p. 847)

In practice, such clear-cut criteria for a recovered memory are rare. Instead, such claims more often involve complex and uncertain processes of remembering and greater subtlety in the complainant’s claims of how and what he or she remembered. Heated controversy still surrounds the debate, with many experts treating cases as if the process of remembering involves either complete fabrication or unequivocal fact. In doing so, many experts retain a firmly entrenched perspective either in favor of a “true, post trauma global amnesia followed by spontaneous or gradual full remembering” (and therefore true) or a false, iatrogenic process of
recovery (and therefore false) argument. Our opinion is that these views emerge from an incomplete consideration of the way in which complainants often claim to have remembered and thus reflect a lack of consideration of the details of each individual case. This can result in an oversimplified presentation of a case in court, in which an expert fails to give due consideration to the range of factors and subtleties that should inform the decision to admit the evidence.

This article is a summary and synthesis of the most relevant psychological and legal literature concerning the controversial and polarized memory recovery debate, and although we must acknowledge the difficulties inherent in attempting to take a neutral stance, we have tried to maintain neutrality. We highlight the notion that in practice, actual cases rarely provide clear-cut examples of either full global amnesia followed by spontaneous recovery or well-remembered accounts. We then move on to a discussion of the legal issues surrounding such cases and highlight how both sides may attempt to argue for an always remembered, recovered, or false view in order to assist their individual cases. We conclude with an analysis of how courts can approach the admissibility and use of recovered memory testimony and related expert evidence in criminal trials. Our central argument is that although experts can helpfully inform the courts as to processes involved in remembering, they should consider several key points that relate to the individual subtleties of the case before providing an assessment of the credibility of an individual specific account. This argument for a case-specific focus, along-side the application of research-based information, reflects arguments made elsewhere for a focus on a pragmatic framework for forensic psychologists (Alison, West, & Goodwill, 2005).

Theories of Repression and Dissociation and Victims’ Abilities to Recall Child Sexual Abuse (CSA)

The concept of repression is deeply rooted in psychodynamic theory. Breuer and Freud (1895/1974) considered repression “a question of things which the patient wished to forget, and therefore intentionally repressed from his conscious thought and inhibited and suppressed” (p. 61). Freud argued that some experiences are so traumatic, such as CSA, that they are buried in the unconscious mind so that the fear and pain associated with them is prevented from overwhelming the victim, although the experience still might have a more limited deleterious effect (see Gampel, 1998). However, he appeared uncertain as to whether the mechanism of repression automatically occurred during the traumatic event itself, meaning that forgetting could be classified as a result of a conscious failure to encode (primary repression), or whether it occurred after the experience, meaning unconscious processes suppress the traumatic memory, a form of motivated forgetting designed to protect the individual from further harm by consciously recollecting the event (repression proper).

These two alternative possibilities have important implications for later recall. Because primary repression is conceptualized as a result of encoding problems, the implication is that a memory of events will not be available to be recovered later because the memory simply was not encoded in the first place. However, because repression proper or motivated forgetting is thought to be an unconscious attempt to avoid thinking about the experience, the implication is that the memory
might be available for recall later on. Janet (1919) stated that avoidance of the anxiety-provoking memory itself would hinder its synthesis and integration at a conscious conceptual level. The experience might not be forgotten; it is simply not thought about to avoid the fear and distress associated with it. This mechanism does not involve any unconscious process; rather the individual chooses not to report the memory for many years for reasons such as embarrassment or fear of punishment.

Another widely reported but controversial mechanism of forgetting of traumatic experiences is dissociation or dissociative amnesia. Brewin (1998) stated that “‘dissociative amnesia’ implies a disturbance in the original encoding of a traumatic event due to an impairment or alteration in consciousness” (p. 216). As individuals are thought to be able to divide their conscious state from the event surrounding them, they are purportedly able to prevent themselves from fully encoding the event. Therefore, when the memory is recovered it might be in the form of flashbacks that contain emotional and sensory components rather than explicit verbal narratives.

In sum, mechanisms have been hypothesized that suggest that memory for CSA could be at some points unavailable to an individual but recovered later on. We now turn to the evidence for the existence of repression and dissociation.

The Difficulty of Interpreting a Complainant’s Account Regarding the Process of Remembering

Several researchers assert that they have established supporting evidence for repression and dissociation through many examples of partial or complete amnesia for traumatic events. For example, Herman and Schatzow (1987) found that 63% of 53 adults in a treatment program for victims of CSA claimed to have experienced partial or complete amnesia, with 74% of the amnesic sample reporting corroborative evidence of the abuse. Feldman-Summers and Pope (1994) found that 40.5% of 79 adults who had been sexually or physically abused as children experienced periods of amnesia. Cossins (1997) noted, in reference to Feldman-Summers and Pope’s study, that “the rates of corroboration for abuse memories are unrelated to whether there had ever been a period of forgetting” (p. 11). However, they used a small and potentially biased sample, and there is the ongoing problem of generalizing from a clinical or community sample to a forensic sample (Connolly & Read, 2003; Epstein & Bottoms, 2002).

This study, as well as many others, presents a persistent problem with retrospective studies because “normal” memory processes such as changes in context, for example meeting a previous abuser or simply thinking about abuse, could potentially cue a previously forgotten memory (Brewin, Dalgleish, & Joseph, 1996; Shobe & Schooler, 2001). In these instances, repression or dissociation is not required to explain memory recovery. For example, Melchert’s (1996) questionnaire-based study of 41 college students revealed that of those that reported varying degrees of amnesia for physical, sexual, or emotional abuse, the majority had made concerted efforts to avoid remembering the events. Only a minority reported being completely amnesic and unaware of the memories, perhaps because individuals choose to wait until adulthood to disclose their
memory of abuse so that they have the necessary coping strategies; resources, both emotionally and financially; and feelings of security, love, and safety.

There is evidence that extended retrieval attempts can lead individuals to believe that at a previous period of their life, they were more amnesic than they in fact were. With increasing effort invested in remembering comes increased overestimation of previous forgetting. Read and Lindsay (2000) found in their study that retrieval efforts could bias retrospective judgments about autobiographical memory. The authors examined autobiographical memories for nontraumatic but, nevertheless, potentially consequential (i.e., significant unusual events that are likely to prove memorable) childhood events such as summer camps and graduations. The participants were asked whether there was a period of time in which they had “less or no memory” for the event. Prior to any sustained attempts to retrieve, 16% claimed less memory, and 5% claimed no memory for specific events. After prolonged retrieval attempts, the rate of perceived partial amnesia rose from 16% to 70% (see also Belli, Winkielman, Read, Schwarz, & Lynn, 1998). That extended retrieval attempts can bias judgments of amnesia raises serious questions about the validity of retrospective studies (Read & Lindsay, 2000).

A number of other studies also suggest that individuals appear to have a poor understanding of the way in which they remember events. For example, Parks (1999) found that individuals claim to forget recalling a childhood event that they thought about only minutes earlier (Joslyn, Loftus, McNoughton, & Powers, 2001). Further, Schooler, Bendiksen, and Ambadar (1997) established that two of their four interviewees claimed complete amnesia for abuse prior to disclosure despite the fact they had previously told others about the abuse during the period of claimed “amnesia.”

Williams's (1995) study is one of the few prospective studies to have used a community sample of 129 adults with a history of CSA. Williams interviewed participants several years later but failed to directly ask about incidents of abuse. Of those that mentioned the abuse, 16% reported a period of amnesia in which they had less memory of the event, and when they were asked to give details, their reports reflected an extremely accurate account of the original details of the documented abuse. Though frequently cited as evidence that memories can be recovered accurately after a period of amnesia, the study suffers from similar reporting problems as highlighted in Read and Lindsay (2000) as well as the associated problem of not asking direct questions about the abuse.

In sum, a central problem in cases involving potentially recovered memories is interpreting the complainant’s explanation for how the abuse was remembered. It is clearly not uncommon for an individual to report that he or she has not had a continuous memory for CSA. People who make these claims can include individuals that have previously always remembered abuse but did not want to talk about it, those who previously remembered abuse but forgot that they were previously aware of abuse, and those who have recovered memories of abuse of which they were previously unaware. The issue of therapy and memory work is also relevant, and we now turn to it.
The Difficulty Associated With Assessing the Origins of “Suggestions” of Abuse

The issue of recovered memories as a product of therapy and potentially inappropriate memory work techniques is controversial, and of course, a persistent problem is that as described in the previous section, we cannot always rely on the memory reports of those remembering abuse or, if applicable, their therapists. Several researchers have argued that memories recovered after long periods of apparent forgetting can emerge in response to poorly conducted therapeutic interventions that either knowingly or unknowingly employ suggestive techniques (Loftus, 1993; Ofshe & Watters, 1994). However, there is some evidence that this is not necessarily the case. Kristiansen (1996) found that 84% of those in that study who recovered memories stated that “the therapist had never even asked if they had an abuse history” (as cited in Cossins, 1997, p. 21). In addition, 40% recovered memories outside of therapy. In Elliott and Briere’s (1995) study, 22% of 116 participants reported partial amnesia, and 20% reported complete amnesia for CSA. Only 8% of those with recovered memories had undergone therapy. This is consistent with the suggestion that normal memory processes may explain at least some aspects of recovered memories.

However, one particular concern involves the overrepresentation of recovered memories among clients of particular therapists (Van Koppen & Crombag, 1999). For example, clients are more likely to recover what many would consider to be implausible memories of alien abduction if they see therapists who strongly believe in alien abduction (Spanos, Burgess, & Burgess, 1994). Similarly, there is an overrepresentation of recovered memories of CSA among psychiatric patients (Orr, 1999). It is unclear whether the latter pattern reflects a causal relationship between abuse and mental disorders or whether such illness makes this population particularly vulnerable to suggestion.

A variety of processes have been thought to influence the production of false memories. Coercion or compliance can occur in instances in which the therapist is perceived as an authoritative figure. Research indicates that clients will admit to claiming pseudomemories they knew were false in order to comply with an experimenter (Barnier & McConkey, 1992). Similarly, Hoelscher, Rosenthal, and Lichstein (1986) established that patients sometimes inaccurately “recall” their behaviors in order to conform and comply with the views and beliefs of their therapists. This reveals that compliance can occur in clinical settings as well as under laboratory conditions. Spanos, Burgess, Burgess, Samuels, and Blois (1999) have also discovered that merely having an authoritative person suggest a certain body sensation (in this case, umbilical itching) is sufficient to generate such memories. Thirty-two percent of age-regression-hypnotized participants and 38% of participants in the control group reported experiencing umbilical itching. Nevertheless, these reports must be treated with some caution as there is some evidence that participants may be simply complying with the demand characteristics of the situation and do not genuinely believe these memories (Spanos, 1992).

Procedures used by certain therapists may be important determinants of generated memories (Bottoms, Shaver, & Goodman, 1991; Mulhern, 1991). For example, Loftus and Pickrell (1995) asked participants to provide imaginary
details of events that had occurred in childhood. The research team provided the to-be-imagined event, in this case, being lost in a shopping mall. Participants gave detailed, confident descriptions of what they “remembered” about an event that never actually happened and that they had been provided with. Similarly, Loftus and Coan (1994) led 5 participants to believe they were lost in a shopping mall when they were 5 years old. Four of the participants recounted very detailed narratives of this imaginary event and refused to believe it was false when they were reminded of the imaginary nature of the experiment. Nevertheless, the fact that all the “ experimenters” were elder siblings making suggestions to younger siblings may have confounded the results. Nevertheless, other researchers have found similar results (e.g., Ost, Vrij, Costall, & Bull, 2002). However, some have criticized these studies on the grounds that the “false” memories could have been imported from actual memories and were merely distortions of those genuine events (Conte, 1999). Pezdek, Finger, and Hodge (1997) demonstrated this by using less plausible events unlikely to have occurred in the participant’s childhood and found significantly fewer false memories. In their first experiment (in which the participants were either Jewish or Catholic), the false possible scenarios were Communion or Jewish prayer. The event was considered less plausible if it was incongruent with their respective religious ideology. Although all 10 accepted the plausible suggestion, only 1 accepted the implausible event. Pezdek et al. concluded that only plausible events could be implanted.

Repetitive questioning can also influence individuals’ readiness to accept prior amnesia for events. As mentioned previously, extended efforts to remember give the illusion of prior amnesia (Read & Lindsay, 2000), and repeated retrieval techniques can also enhance individuals’ commitment to errors. In summary, Lindsay and Read (1995) concluded that the perceived authority and trustworthiness of the source of suggestion, repetition of the suggestion, the plausibility of the event being suggested, imagination, and lowering of memory-monitoring response criteria all contribute to producing memories that may be either distortions or complete fabrications that are subsequently held with great conviction by those who have remembered such events. DelMonte (2000) suggested that if “psychotherapists would follow recommended clinical practice, for example, by not being forcibly directive, avoiding strongly suggestive comments, monitoring their own counter-transferences, and not imposing ideologically motivated ‘explanations’ and ‘interpretations,’ some of this confusion might abate” (p. 10).

In sum, procedures that rely on suggestion can create inaccurate memory reports, and these procedures might be especially likely to appear within therapy. Therefore, it is important that the courts should consider potential “interference” from therapists carefully, although it is neither a necessary nor a sufficient condition for false accounts to be produced. Nevertheless, the literature that has been reviewed so far does not focus on the potentially unique impact of CSA on memory. We now turn to this issue.

The Difficulty of Generalizing From Empirical Research to Cases of CSA

A fundamental question exists concerning whether trauma enhances or inhibits memory. The trauma superiority argument contends that trauma enhances
memory for events rather than inhibits it, especially the central details (e.g., Shobe & Kihlstrom, 1997; Wagenaar & Groenweg, 1990; Yuille & Cutshall, 1986). In contrast, the **traumatic memory argument** asserts that traumatic events result in different encoding, storage, and retrieval processes and can easily be forgotten for long periods until particular cues facilitate retrieval (Van der Kolk & Fisler, 1995; Van der Kolk, Hopper, & Osterman, 2001).

Research from human and animal studies tends to suggest that trauma and stress can enhance the memory of the experience. Berntsen (2001) explained how the brain releases a stress hormone that aids in the consolidation of memory during high arousal states. Indeed, Alvarez (1992) asserted that a common problem for trauma survivors is an inability to forget the event. As long ago as 1890, James claimed that a highly traumatic event could be so stressful and emotional that it would almost “leave a scar upon the cerebral tissues” (as cited in Porter & Birt, 2001, p. 102). Several studies have supported this basic principle. For example, Yuille and Cutshall (1986) found memory to be intact for witnesses to murder; Wagenaar and Groenweg (1990) found that concentration camp survivors retained intact memories of their experiences; and Porter and Birt (2001) found that the individuals in their study had intrusive and repetitive memories that were rich, coherent, and detailed.

However, Chapman and Underwood (2000) suggested that memories of a traumatic event vary depending on the level of stress associated with the event. They found that moderately stressful events, such as a near accident in a car, resulted in memory impairment, with 80% of incidents forgotten within 2 weeks. However, higher levels of arousal, involving actual collisions rather than near misses, led to detailed memories. Of course, there are considerable differences between sexual abuse and car accidents and near accidents, and participants would have had less of a reason to rehearse a near accident than an actual accident. Joslyn, Carlin, and Loftus (1997) established a strong positive relationship between comprehensive understanding of an event and its perceived significance, as well as a positive relationship between the number of self-reported incidents of thinking about an event and the probability of being able to accurately recall it.

Easterbrook (1959) argued that trauma can both inhibit and enhance memory depending on the level of arousal and stress involved, with moderate arousal enhancing memory but extreme arousal causing interference with encoding due to a narrowing of attention (Byrne, Hyman, & Scott, 2001). However, Shobe and Kihlstrom (1997) pointed out that there are no laboratory studies to support the hypothesis that central details of an event can be entirely forgotten.

Van der Kolk and Fisler (1995) argued that in order to demonstrate the “special” nature of traumatic memories, studies need to measure the characteristics and content of traumatic memories over time and in comparison to nontraumatic memories. They asserted that the many experimental studies of memory are of little significance because such studies do not involve highly stressful and traumatic stimuli. Van der Kolk and Fisler (1995) stated,

If trauma is defined as the experience of an inescapable stressful event that overwhelms one’s existing coping mechanisms, it is questionable whether findings of memory distortions in normal subjects exposed to videotaped stresses in the
laboratory can serve as meaningful guides to understanding traumatic memory. (p. 506)

But of course the difficulty here lies in what is ethically permissible in the laboratory.

Given that ethics bar researchers from testing the hypothesis that suggestions can give rise to false memories of CSA, there will always be room to argue that studies demonstrating false memories of other kinds cannot be generalized to memories of CSA. (Lindsay & Read, 1995, p. 867)

Using the Traumatic Memory Inventory procedure, Van der Kolk and Fisler (1995) and Van der Kolk, Burbridge, and Suzuki (1997) found that in contrast to nontraumatic memories, traumatic experiences are initially retrieved as loosely connected chunks of sensory information and flashbacks in a fragmentary and confusing form. The sensations include visual images, smells, sounds, affective states, and bodily sensations that are associated with the traumatic experience intruding into consciousness. Van der Kolk et al. (2001) argued that these fragments of sensory information represent implicit memories of the experiences encoded in place of explicit, narrative memories. The latter are less available to consciousness because of the individual’s dissociative state at encoding. The suggestion is that the stress of the event interferes with consolidation of explicit memories. However, the inhibition of explicit memory formation fails to influence implicit memory. Over the course of a professionally conducted therapeutic intervention, they claim the flashbacks and fragmentary pieces of information can be constructed into a verbal account. Van der Kolk and Fisler’s (1995) study has been criticized more recently by Gray and Lombardo (2001), who observed that there was no control group and that the traumatic and nontraumatic events chosen were not matched for age of occurrence, with the traumatic memories generally being from childhood and the nontraumatic memories from adulthood. Therefore the differences found could have been due to conventional processes of decay of the childhood trauma or due to infantile amnesia. Further, the advertisement for participants required individuals who were “haunted” by a traumatic memory. This may have selectively biased the participant group by discouraging individuals with explicit memories of traumatic experience (Shobe & Kihlstrom, 1997). Finally, the study failed to corroborate individuals’ accounts.

A neurological justification has also been suggested by Van der Kolk et al. (2001) for differences in traumatic and nontraumatic memories. They found that positron-emission tomography scans of people with posttraumatic stress disorder revealed increased activation in the right hemisphere during a traumatic memory, an area considered dominant in evaluating the emotional significance of sensory information. They proposed that narrative verbal accounts of traumatic memories are difficult to recall because the hippocampal memory system fails under extreme stress. Van der Kolk et al. (2001) concluded that “both interviews and brain imaging of traumatized people confirm that traumatic memories come back as emotional and sensory states, with limited capacity for verbal representation” (p. 28).

Alongside Van der Kolk et al.’s (2001) argument that trauma memories are distinct from conventional memories is the more recent proposal that CSA
memories are qualitatively different from other trauma memories. A study by Mechanic, Resick, and Griffin (1998) appears to support Van der Kolk and Fisler (1995) and Van der Kolk et al.'s (1997) work in establishing that the memories of people who were raped improved over time, with 37% reporting significant amnesia for the experience 2 weeks afterward, and 16% reporting similar levels 3 months after. Roe and Schwartz (1996) established that 60% of the clients in their sample reported initially recovering memories of abuse in the form of flashbacks and sensations. Clients were able to put the traumatic experience into a continuous verbal narrative only after time. In a study comparing individuals who had experienced amnesia for abuse with individuals with continuous memories of the abuse, Cameron (1996) established that the former were more likely to report sensory and fragmented memories and less able to articulate the experience. In Burgess, Hartman, and Baker’s (1995) prospective study of 34 children who had been abused, they established that although at the time clients had both implicit and narrative memories for the abuse, 5 to 10 years later many had lost the narrative aspects of the memory but retained implicit aspects of the memory (i.e., flashbacks). Many authors have cited these studies as evidence of dissociation, despite the many limitations of self-report measures (see the arguments outlined in the opening section of this article). However, a key feature of all these studies is that they include clients’ efforts to remember the event, which in itself may encourage the creation of a narrative.

More recently, Gray and Lombardo (2001) failed to find support for Van der Kolk and Fisler’s (1995) or Van der Kolk et al.’s (2001) view that traumatic memories are “special.” In Gray and Lombardo’s work, nontraumatic memories also occurred initially as sensations and subsequently in narrative form. Thus, nontraumatic memories evolved in much the same way as memories for traumatic events, with fragmentation and disorganization in the early stages and more detailed narratives emerging over time. Koss, Figueredo, Bell, Tharan, and Tromp (1996) suggested that the reason memories evolve in this way is due to lack of rehearsal at the early stages, with increasing rehearsal and consolidation over time.

Similarly, Porter and Birt (2001) found that traumatic and nontraumatic memories were comparable in vividness, coherence, and overall quality even though clients spent more time ruminating over nontraumatic events. Berntsen (2001) has established that involuntary memories involving emotional, behavioral, and physiological reliving are not just limited to traumatic experiences. Berntsen found that flashbacks could occur for “peak” events as well as traumatic ones. Further, Read and Lindsay (2000) found that nontraumatic memories could be forgotten and triggered in the same way as traumatic ones, and Shobe and Kihlstrom (1997) argued that the evidence presented for traumatic memories being “special” is anecdotal and comes from clinical evidence drawn from poorly controlled and confounded investigations.

In summary, an extensive laboratory literature suggests that traumatic memories are likely to be well remembered, particularly the central details. However, with regard to CSA, ethics makes the exploration of accurate and suggested memories impossible under laboratory conditions. Self-report data do suggest that memories for actual events can gradually return in a similar manner to the way that recovered memories are reported to be retrieved.
Summary of the Psychological Issues

In reviewing the psychological evidence, it is clear that considerable debate persists with regard to the proposed processes of recovering memories. Several researchers claim evidence for a special and discrete mechanism for forgetting and recovering traumatic memories. Others argue that conventional memory processes can account for this process, whereas still others argue that the contemporary view of memory presents compelling evidence that such events cannot be easily forgotten. In addition to presenting the diversity of opinion, we have argued that there are several other difficulties in assessing the credibility of any given account. These include the difficulty of establishing how the complainant claims to have remembered, the difficulty of establishing the extent to which other individuals may have shaped the account, and the difficulty of applying general research findings to specific cases, especially concerning CSA.

In legal proceedings, the interest is in individual cases, not in what is likely at a group level. Therefore, in the next section we examine the legal difficulties associated with providing such information in reference to the credibility of a complainant’s account. In addition to discussing the issues surrounding the lack of agreement in the scientific community, we outline further problems in classifying a case, the admissibility of such evidence, and the possible options available to the courts in dealing with such cases.

The Legal Effects of Recovered Memories in Criminal Cases

When CSA is prosecuted many years after the alleged events, courts typically distinguish between two categories of cases (Johnson v. Johnson, 1988). In what can be called Category 1 cases, the complainant has always remembered the abuse but has been psychologically unable to complain because of feelings of betrayal, powerlessness, or guilt. In so-called Category 2 cases, the complainant has failed to complain because of forgetting the memories of abuse, which have subsequently been recovered. As previously outlined, there is evidence that at least some Category 2 cases involve false memories. It is therefore of crucial importance that the criminal justice system be able to sort out those claims that are based on true recovered memories from those that are not.

Although this article concentrates on Category 2 cases, we cannot assume that cases of always-remembered abuse (Category 1) and cases of recovered memories (Category 2) are distinct and distinguishable—in other words, that a complainant will fall within either Category 1 or Category 2 and that it is possible to identify the category into which a complainant falls. Such an assumption may be invalid in some cases, as participants may intentionally or unintentionally attempt to shift the case from one category to the other, or the case may fall within both categories. For example, if cases involving recovered memories of abuse are less reliable than those in which the complainant has always known of the abuse but has delayed reporting, defendants may attempt to characterize always-remem-

---

1For a more detailed analysis of the legal arguments across a number of different jurisdictions, the interested reader may wish to consult Chapter 7 of Delayed Prosecution for Childhood Sexual Abuse (Lewis, 2006b, pp. 153–177), from which this section is adapted by permission of Oxford University Press (www.oup.com).
bered abuse as falsely recovered memories (see, e.g., *R. v. JP.*, 1999). Conversely, in order to avoid the difficulties surrounding recovered memories, complainants who have recovered memories may characterize their experience as falling within Category 1 rather than Category 2. This could be done intentionally following advice from a therapist or could be the effect of suggestion or confabulation during therapy if a complainant comes to believe, incorrectly, that he or she has always remembered the abuse. Potentially, a case may also fall within both categories if some memories of abuse were always remembered and some have been repressed and recovered (*E. v. R.*, 1997). As far as these cases are concerned, those memories that were recovered should be subject to the approach we propose below in relation to Category 2 cases.

When the complainant has presented the case as falling within Category 1 (always-remembered abuse) but the defense alleges it is a Category 2 case, an analogy to the recently evolving judicial approach to the testimony of sexual assault complainants is apposite. Warnings to the jury regarding the need for caution in assessing such testimony are now given only when there is an evidentiary basis to support the claim that the witness is unreliable (Lewis, 1996, 2006a). Thus, only when there is an evidentiary basis to support the defense claim of unreliability should the case be treated as falling within Category 2. Such an evidentiary basis would include therapist’s notes indicating that memories have been recovered or prior inconsistent statements by the complainant contradicting her or his account that she or he has always remembered the abuse.

The mere allegation by the defense that the case falls within Category 2 should not be sufficient to warrant Category 2 treatment; otherwise all delayed prosecutions of CSA could become subject to this treatment, despite the absence of serious concerns about the reliability of testimony in cases that do not involve recovered memories. If the defense simply alleges that the complainant is lying when she or he claims that she or he has always remembered the alleged abuse (without an evidentiary basis to support this allegation), then it is an issue of credibility that a jury is equipped to assess (see, e.g., *R. v. Thorne*, 1995).

After noting that the categories identified above may be both malleable and manipulated, we focus our attention on how courts should approach the admissibility and use of recovered memory testimony and expert evidence in criminal trials. There are several existing and potential legal tools that can influence the pursuit of prosecutions for CSA based on recovered memories (for a more detailed discussion of international law, see Lewis, 2006b).

**Stopping the Case Before Trial**

Cases may be ended in two specific ways—by the police simply deciding that no crime has occurred and declining to pass information along to prosecutors or by prosecutors deciding not to prosecute (Sanders & Young, 2002). In addition, different jurisdictions have varied means by which a defendant may attempt to stop a long-delayed case that is based on recovered memories. However, in jurisdictions that have a statute of limitations for serious criminal offenses, such as many U.S. states, the constitutionally entrenched human rights protection of the right to a speedy trial is generally inapplicable to the delay between the alleged offense and charge (Lewis, 2006b; Lewis & Mullis, 1999).
Trial-Related Issues

Before we discuss the criminal courts’ responses to recovered memory cases, the legal situation with regard to the admissibility of hypnotically refreshed testimony is worth considering. These cases are perhaps the most analogous to Category 2 cases. Forensic or investigative hypnosis has been used in various jurisdictions to retrieve memories from witnesses (Gudjonsson, 2003). In the late 1960s, U.S. courts initially welcomed hypnotically refreshed testimony, and the use of hypnotism to refresh memory was considered comparable to the use of memory refreshing documents (Kanovitz, 1992). In *Harding v. State* (1968), the Maryland Court of Appeal ruled that the issue of hypnotic refreshing of a witness’s memory was relevant to credibility and weight rather than admissibility. This position was followed by numerous U.S. courts until the late 1970s, when serious concerns were raised in the academic literature about the reliability of hypnotically refreshed testimony (Diamond, 1980; Orne, 1979) on the basis of the dangers of confabulation, memory hardening, suggestion, and conflation (Kanovitz, 1992; Kebbell & Wagstaff, 1999).

Subsequent courts (and legislatures) diverged in their approaches (see Eisenberg, 1995, for a review). A minority of U.S. states continued to admit the testimony. Others imposed stringent procedural safeguards governing the hypnosis. If these safeguards were not observed, witnesses who had undergone hypnosis were not permitted to testify regarding their refreshed memories. Some U.S. courts (including the federal courts) mandated a slightly less rigorous, more holistic *totality of the circumstances* pretrial review to ensure the reliability of the testimony under which compliance with safeguards is a factor rather than determinative. In recent years, some academics have argued against wholesale exclusion of hypnotically refreshed testimony in favor of individual reliability assessments (e.g., Brown, Scheflin, & Hammond, 1998).

These experiences with hypnotically refreshed testimony suggest a spectrum of possible legal–policy approaches to repressed memory testimony. First, using the law of expert evidence, all such testimony could be excluded as the product of one or more unreliable techniques, or some subset could be identified for exclusion. Second, the admission or exclusion of the testimony could be determined on a case-by-case basis either before the trial begins or at the commencement of the trial by using a voir dire at which the reliability of the testimony would be assessed. Third, such testimony could be admitted but considered inherently suspect. The jury could receive appropriate warnings, and the trial judge could be required to direct a verdict of not guilty if the complainant’s testimony is weak and no supporting evidence exists. Fourth, such testimony could simply be admitted, with cross-examination, appropriate expert evidence, and jury instructions forming safeguards. The first, second, and fourth options mirror the judicial approaches to hypnotically refreshed testimony outlined above, and we now discuss each option in turn.

**Option 1: Automatic Exclusion**

One method for achieving automatic exclusion would be to use the law of expert evidence. In two consolidated New Hampshire cases involving indictments for aggravated felonious sexual assault, the trial judge held that “testimony that is
dependent upon recovery of a repressed memory through therapy cannot be logically disassociated from the underlying scientific technique” (State v. Hungerford, 1995, p. 1). The Supreme Court of New Hampshire agreed with the trial court that “a recovered memory that previously had been completely absent from a witness’s conscious recollection . . . cannot be separated from the process, if any, that facilitated the recovery” (State v. Hungerford, 1997, p. 921). Following a pretrial hearing on the psychological evidence of repressed memories, the trial judge ruled that the complainant’s testimony was inadmissible as the State had “failed to meet its burden of proving that there was general acceptance of the phenomenon of repressed memories in the psychological community, and, further, that the State had failed to demonstrate that the phenomenon was reliable” (State v. Hungerford, 1997, pp. 919–920).

Some courts and commentators have questioned whether the law of expert evidence can be used to exclude eyewitness testimony perceived as unreliable: to . . . argue that the standard [for admission of expert evidence] controls the admission of the testimony of an ordinary witness for which the only criterion traditionally has been only personal knowledge of the relevant information is somewhat unprecedented and subject to debate. (Zoltek-Jick, 1997, p. 470)

Similarly, Eisenberg (1995) described the exclusion of such testimony as “improper” (p. 268), relying on the Daubert (1993) court’s recognition that “there are no certainties in science” (p. 2795). Some precedent for this unusual use of the law of expert evidence does exist in the hypnosis cases, of which a good example is People v. Guerra (1984). The use of expert evidence to exclude recovered memory testimony has now been accepted or advocated by a number of U.S. courts (e.g., Commonwealth v. Crawford, 1996; Franklin v. Stevenson, 1999; State v. Hungerford, 1997; State v. Quattrocchi, 1996).

Option 2: Pretrial or Voir Dire Reliability Assessment

The second possible approach to the admissibility of recovered memory testimony would involve an assessment of the evidence in each case either before the trial commences or at its outset. This could be accomplished in one of two ways. The first would be a case-by-case version of the approach outlined above, linking the admission of recovered memory testimony with the reliability of the psychological techniques used to recover the memories. This less absolute approach was proposed by the appeal court in the Hungerford case, in which it was stated that “testimony that relies on memories which previously have been partially or fully repressed must satisfy a pretrial reliability determination” (State v. Hungerford, 1997, p. 921). The relevant reliability comparison is with the accuracy of ordinary memory, rather than with the truth of the recovered memory. The Supreme Court of New Hampshire enumerated a number of relevant considerations:

In determining the reliability of a recovered memory,—that is, whether the recovered memory is reasonably likely to be as accurate as ordinary memory—the trial court should consider the following factors: (1) the level of peer review and publication on the phenomenon of repression and recovery of memories; (2) whether the phenomenon has been generally accepted in the psychological com-
munity; (3) whether the phenomenon may be and has been empirically tested; (4) the potential or known rate of recovered memories that are false; (5) the age of the witness at the time the event or events occurred; (6) the length of time between the event and the recovery of the memory; (7) the presence or absence of objective, verifiable corroborative evidence of the event; and (8) the circumstances attendant to the witness’s recovery of the memory, i.e., whether the witness was engaged in therapy or some other process seeking to recover memories or likely to result in recovered memories. (State v. Hungerford, 1997, p. 925. Citations are omitted.)

The first four of these factors are from Daubert (1993). There is substantial disagreement among U.S. courts as to whether the phenomenon of recovered memories passes these four tests from Daubert, with some courts expressing doubts as to reliability on the basis of Factors 2, 3, and 4 (State v. Hungerford, 1997; State v. Quattrocchi, 1996). Unsurprisingly, criminal courts seem more doubtful than civil ones. According to the court in the Hungerford case, in which the memories had been recovered through therapy, further inquiry would be needed, including

an examination of the therapist’s qualifications, the type of therapeutic approach used, whether complaints of false accusations have been filed against the therapist, whether the therapist ordinarily seeks hidden memories or believes that many psychological problems stem from sexual abuse, and whether the therapist remains detached during the process or “validates” allegations of abuse that arise. (State v. Hungerford, 1997, p. 925. Citations are omitted.)

This more individualized approach could also apply to spontaneously recovered memories, despite the lack of a link to the expert evidence via the use of scientific techniques to recover the complainant’s memories. A few months after the appellate decision in State v. Hungerford (1997), the New Hampshire Supreme Court applied the same approach to a case in which the complainant’s memories had been spontaneously recovered outside of therapy and held them unreliable and inadmissible in light of Factors 5, 6, and 7 of the Hungerford test and the fact that the complainant’s memory recovery had apparently occurred during dreams (State v. Walters, 1997). The basis for this decision is unclear because the Hungerford courts’ use of expert evidence to exclude recovered memory testimony was explicitly linked to the scientific techniques used to recover the memories (State v. Hungerford, 1995; State v. Hungerford, 1997). The wide approach suggested in State v. Quattrocchi (2001), relying on the necessity of expert evidence to assist the jury, would provide one possible link here.

The second method of individualized pretrial assessment would be to use the trial judge’s discretion(s) to exclude evidence. This type of approach has also been proposed in the United States under Federal Rule of Evidence 403. Rule 403 provides that

although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.

Applying this rule to recovered memory, Smith-Lee (1996) has argued that
the probative value of recovered memory testimony in general, and a given witness’s testimony in particular, is low because of the possibility of suggestion and confabulation, and because of the uncertainty in the mental health profession about the veracity of recovered memories. The prejudicial effect, on the other hand, might be high. The mere accusation of sexual abuse is devastating, and could prejudice the jury against a defendant even before any evidence is offered. In the case of recovered memory, cases are brought years, even decades, after the alleged event occurred, frequently leaving the recovered memory witness’s testimony as the only “hard” evidence available. Finally, the sentiment expressed by clinical psychologists about their patients’ memories may be common among jurors as well: the idea that nobody would invent such a horrible story in the absence of some truth. (pp. 636–637)

**Option 3: The Suspect Evidence Approach**

In a case that depends wholly or substantially on repressed memory testimony that the defendant alleges is false, the trial judge could be required to warn the jury of the special need for caution before convicting the defendant in reliance on the testimony. Further, the trial judge could instruct the jury as to the reasons for the need for such a warning and refer to the possibility that a mistaken witness can be a convincing one and to the danger of memory hardening. In cases in which the repressed memory testimony is poor or weak, the trial judge should withdraw the case from the jury and direct an acquittal unless there is other evidence that supports the testimony. Such supporting evidence would not have to amount to corroboration in the technical sense. The supporting evidence “may be corroboration in the sense lawyers use that word; but it need not be so if its effect is to make the jury sure that there has been no mistaken identification” (R. v. Turnbull, 1977, p. 230).

Such an approach is not without precedent; some state legislatures and state civil courts require corroboration of alleged abuse before tolling (suspending) or extending the statute of limitations (e.g., Moriarty v. Garden Sanctuary Church of God, 2000; Olsen v. Hooley, 1993; Petersen v. Bruen, 1990; Okla. Stat. Ann. tit. 12 § 95[6], 2005). Perhaps an even closer analogy in the context of the admission of hypnotically refreshed testimony is found in two decisions by U.S. federal circuit courts of appeal. In United States v. Valdez (1984), the Fifth Circuit required corroboration before admitting hypnotically refreshed testimony, whereas in Sprynczynatyk v. General Motors Corp. (1985), the Eighth Circuit used the presence of corroboration as a factor when determining the admissibility of such testimony.

The central issue concerning this option is the question of when testimony of recovered memories would be considered weak. Arguably, the questions raised about the reliability of recovered memories suggest that all such testimony should be considered weak and therefore subject to the supporting evidence requirement. According to Brandon, Boakes, Glaser, and Green (1998), “there is no means of determining the factual truth or falsity of a recovered memory other than through external evidence, difficult though this is to obtain” (p. 304). Similarly, Haber and Haber (1998) pointed out that the most reliable method of distinguishing between true and false recovered memories is through the use of supporting or corroborative evidence. Alternatively, because memories recovered through therapy are
subject to greater concerns regarding suggestibility, conflation, and confabulation, only those memories could be considered weak. However, until these concerns can safely be said to be minimal in the case of nontherapeutically recovered memories, in comparison to ordinary eyewitness testimony, all recovered memory testimony would be considered weak potentially and subject to the supporting evidence requirement.

Unfortunately, given the delay between the alleged offense and the trial, it will be rare that supporting evidence will be available, but the following potential supporting evidence has been suggested and used in cases (Haber & Haber, 1998; Leo, 1997; Lewis, 2006b; Lewis & Mullis, 1999):

1. A pretrial admission or confession by the defendant or an admission during his or her testimony (although care is warranted as there is some evidence of false confessions in this context; e.g., see, Loftus & Ketcham, 1994).

2. Lies by the defendant, although contradiction of the defendant’s account by the complainant would not be sufficient as the complainant would not be considered to be an independent witness.

3. Admissible evidence of the defendant’s prior or subsequent misconduct or similar allegations by another witness, unless there is a danger of collusion or contamination.

4. Physical forensic or medical evidence.

5. School records.

6. Photographs or recordings of the abuse.

7. Confirmation of other factual events recalled by the complainant.

8. Independent eyewitness accounts.

Option 4: Testimony Admissible

The most liberal approach would be to presume recovered memory complainants to be competent and therefore admit their eyewitness testimony. Any concerns about the reliability of the recovered memories would go to weight rather than to admissibility, consistent with the approach originally taken in the United States toward hypnotically refreshed testimony (Harding v. State, 1968). The trial judge would give a warning to the jury regarding the reliability of the testimony, in addition to the standard delay warning (Lewis, 2005; Lewis & Mullis, 1999). This warning could be bolstered by reference to expert evidence and “careful instruction on the burden of proof” (Daubert, 1993, p. 2798). The defense would rely on expert and other contrary evidence and cross-examination of the complainant to cast doubt on the reliability of his or her testimony. The risks of suggestion, confabulation, conflation, and memory hardening could be explored in cross-examination with the help of expert evidence on the association of these risks with memory recovery.
The Best Approach?

In light of the current absence of scientific consensus on the reliability of recovered memories that has been outlined here and the serious doubts over the credibility of at least some proportion of these memories, in our view the only approach presently consistent with the presumption of innocence is to require supporting evidence in recovered memory cases (Option 3). Thus, in cases in which one is relying on memory evidence alone, we recommend that expert opinion should not be relied on as the sole basis for informing the jury as to the credibility of a complainant’s account.

Option 4 should be rejected because of the danger of miscarriages of justice. The effectiveness of cross-examination of a recovered memory complainant is limited because of the problems of memory hardening (Smith-Lee, 1996). Jurors may be overly impressed by recovered memory testimony, believing it to be more accurate than nonrecovered delayed memory testimony (Coleman, Stevens, & Reeder, 2001; Wagstaff, Vella, & Perfect, 1992). Experts have widely differing views and therefore may not be able to assist the jury with their evaluation of the testimony (Smith-Lee, 1996). Although such controversy exists over the phenomenon of recovered memories, it is difficult to sustain the argument that expert evidence can dispel “the alleged myth that people do not forget their abuse” (Freckelton, 1996, p. 28), particularly as the evidence from mock juror studies is equivocal. One study suggests that expert testimony does not affect trial outcome in civil recovered memory cases (Stewart, Whiteside, & Golding, 2000). Instead, prior beliefs in the validity of recovered memories strongly influenced whether the defendant was found guilty (Griffith, Libkuman, Kazen, & Shafir, 1999; Stewart et al., 2000). Another study provides evidence that jurors are more likely to find the defendant not guilty when the defense offers expert evidence than when it does not (Sugarman & Boney-McCoy, 1997). A third study found that plaintiff expert testimony was influential even if countered by a defense expert (Griffith, Libkuman, & Poole, 1998), although a later study found that varying types of specific and general expert testimony for both plaintiffs and defendants had no effect on verdicts (Griffith, Libkuman, Dodd, Shafir, & Dickinson, 2002). Further, doubts have been cast on the efficacy of directions to the jury, and there is some evidence to suggest that corroboration warnings may be counterproductive (Temkin, 2000).

Option 1 should also be rejected on the grounds of its absolutism. It would apply to all recovered memory cases, regardless of the existence of supporting evidence. In cases with supporting evidence, the jury will be assisted in its evaluation of the complainant’s testimony by the supporting evidence. Automatic exclusion would therefore be an overly cautious approach.

Option 2 requires examination of the evidence prior to trial, thus negating the possibility that unanticipated evidence suggesting reliability may emerge during the trial. Although its increased flexibility to examine recovered memories on an individual basis is an advantage over Option 1, the lack of a clear rule requiring supporting evidence is problematic given the difficulties that will be faced by a trial court in assessing reliability in the absence of supporting evidence.

If we are truly committed to the principle underlying the aphorism that “it is better to let 10 guilty men go free than convict 1 innocent one,” then this
commitment must extend to cases involving recovered memories of CSA. It is, of course, of prime importance to act on complaints of CSA. Complainants in these cases will almost invariably appear to be honest, convincing, and compelling. It is unrealistic to expect judges and juries to assess the reliability of unsupported recovered memories when experts in the field are unable to do so, even though this is likely to reduce the number of convictions of guilty as well as innocent defendants. Adopting this high standard may also shield victims of sexual offenses who have not recovered memories from skepticism that might result from false convictions based on recovered memories and generalized to all complainants.

Conclusion

We have highlighted a number of difficulties for psychologists who become immersed in debates concerning a recovered memory of abuse and provided recommendations for overcoming these hurdles in a courtroom setting. Underlying all of these recommendations is the proposal for a closer focus on a case-based analysis. First, experts should ask detailed questions about the way that individuals remember things. When possible, experts should seek documentary evidence to construct a timeline, for example, using patient notes and witness statements. Second, experts should scrutinize issues that may have influenced memory production. Although it is extremely difficult to obtain a detailed picture of the context within which individual accounts are shaped, it is inappropriate to assume that an individual’s participation in therapy is proof alone that an account is false or even distorted. Similarly, the fact that an individual has not attended therapy does not necessarily reduce the probability that an account may be distorted or false, because research indicates there are many diverse influences beyond iatrogenic processes. As far as is possible, experts should seek to establish when the event is said to have happened, who was told of this information (e.g., when and how often), and what other subsequent (and, potentially, prior) events may have shaped encoding and retrieval. Third, experts should exert caution in relying on general models to discuss individual cases. Because of the substantial variation in individuals’ accounts of the alleged abuse (e.g., severity, delay in reporting, one-off or persistent offenses, the extent to which other influences may have shaped memory, individual variation in what is considered traumatic), it is not appropriate to draw exclusively on arguments for or against delayed accounts without due consideration of these multifarious pathways to reporting the abuse. A key benefit for the courts may lie in experts explaining their views on how these various features (e.g., consequentiality, decay, repetition, social influence, individual differences) influence memory, aiding the jury’s understanding of the individual case.

References


Fed. R. Evid. 403.


People v. Guerra, 690 P.2d 635 (Cal. 1984).


Spynczynatyk v. General Motors Corp., 771 F.2d 1112 (8th Cir. 1985).


United States v. Valdez, 722 F.2d 1196 (5th Cir. 1984).


Received March 20, 2006
Revision received August 21, 2006
Accepted August 23, 2006