Challenging the Eyewitness Expert

by

John C. Yuille, Judith (Cutshall) Daylen, Stephen Porter & David Marxsen

INTRODUCTION

In an attempt to apply their findings to a practical problem, psychologists began to examine eyewitness memory almost one hundred years ago (Yuille et al., 1994). Eyewitness memory researchers have examined two forms of memory testing: recall and recognition. Research concerned with recall has focused on the ability of victims and witnesses to remember and verbally reproduce details about an event. The amount and accuracy of the recalled information has generally been the primary interest of this research. On the other hand, eyewitness recognition research has investigated the proficiency of witnesses in identifying the perpetrator of a crime and the types of errors that witnesses make in their identifications.

Both recognition and recall have been studied in the same basic manner: laboratory-based (controlled) research. The research paradigm typically involves undergraduate students witnessing some event, followed by a memory test involving either recall or recognition, or both. The event witnessed by the students is often a recreated criminal act that is presented via a sequence of slides (e.g., Loftus et al., 1978) or a video tape (e.g., Bregman and McAllister, 1982). Alternatively, in some studies the "criminal" event is staged before unsuspecting students. For example, the students might report to a psychology laboratory for what had been advertised as a study of memory. Unexpectedly, they witness the "theft" of a piece of equipment, and they are then asked to provide a report of the staged theft and/or attempt an identification of the thief from a lineup or photo-spread. More recently eyewitness research has moved out of the laboratory into more naturalistic contexts (elaborated below).

A recurring theme of eyewitness research has been the fallibility and malleability of human memory, whether measured by recall or recognition. Most researchers have focused on delineating the limitations and problems associated with eyewitness performance. The factors investigated have included the impact of delay (between the event and the memory test), the type of memory test employed (e.g., free recall versus leading questions), the form of lineup used, or the type of identification instructions given. In this chapter we provide an overview of the factors that have been examined in contemporary eyewitness research, with a shift in emphasis from the limitations and problems with eyewitnesses to an emphasis on the limitations and problems associated with eyewitness research. The corresponding problems with the validity of expert testimony on eyewitness competence is also a theme of this chapter.

The concerns that we raise in this chapter emerge from a critical examination of contemporary eyewitness research that leads one to question the generalizability of this research. As noted above, the vast majority of these studies have been conducted in laboratories or other contrived situations, even though the context in which the research results are applied is the forensic world: civil, family, and criminal courts; crime and accident scenes; and other events of consequence to victims and witnesses. Although this chapter is chiefly concerned with memory issues for events of a criminal nature (e.g., physical assault), the same challenges we discuss should confront the eyewitness expert in civil courts. The question thus arising: can the results of research conducted in one context, i.e., the laboratory, be applied to another, i.e., the courts? This question provides the basis for many of the challenges to expert testimony concerning eyewitness memory (Elliot, 1993; Konecni and Ebbersen, 1986; McCloskey and Egeth, 1983; Yuille and Cutshall, 1986; Yuille, 1993) that are raised in this chapter.

---

1 John C. Yuille, Ph.D., is a full Professor of Psychology at the University of British Columbia and his work is supported by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC).

Judith (Cutshall) Daylen, Ph.D., is a clinical psychologist in private practice in Vancouver, British Columbia.

Stephen Porter, M.A., is a doctoral candidate in psychology at the University of British Columbia and is supported by a doctoral fellowship from SSHRC.

David Marxsen, B.A., is an M.A. candidate in psychology at the University of British Columbia.
Before a review of contemporary eyewitness research is presented, the reader would benefit from knowing that the debate about the applicability of eyewitness research and the usefulness of expert testimony on eyewitness memory is not a new one. Early in this century a short, but heated, exchange occurred between some psychologists and members of the legal profession that foreshadowed the contemporary situation. We begin with a short overview of this historical debate.

**THE AUSSAGE PERIOD**

"Whenever the Psychologist is ready for the Courts, the Courts are ready for him." (Wigmore, 1909)

This quote is from a scathing review of Hugo Munsterberg’s controversial text, *On the Witness Stand* (1908). J.H. Wigmore, perhaps the leading American legal authority of his day, was responding to the landmark psychological work that summarized the previous decade of research in the new field of eyewitness testimony. The thesis of Munsterberg’s text was that after a decade of fecund *Aussage* research, the time was right for psychologists to bestow upon the legal system their learned commentaries regarding testimonial evidence.

*Aussage* (German for report/testimony) research had its roots in turn-of-the-century Germany. The origin of experimental psychology, traditionally dated from the establishment of the first psychology laboratory by Wilhelm Wundt in 1879 at the University of Leipzig, had only recently occurred. This young field of study was still trying to establish itself as a legitimate science independent of its parent discipline, philosophy (Boring, 1957). *Aussage* research, with its twin emphases on the experimental method and the application of its results to the "real world" (in this case the courts), was in many ways an ideal means for achieving psychology’s goals of becoming a legitimate, scientific discipline. The *Aussage* researchers were typical of the first generation of experimental psychologists in their efforts to build a positivistic science modeled after the physical sciences that would benefit, and be respected by, society. A pioneer French psychologist, Alfred Binet (1900), had argued for the "advantage that would accrue from the creation of a practical science of testimony."

The characteristic *Aussage* study was the so-called "reality experiment" in which an event would be staged in front of a group of people. For instance, a university lecture might be interrupted by an argumentative student (actually a confederate of the experimenter) who drew a gun and fired into the air. The observers would be asked to recount all that they could remember about the incident, usually after they had been informed that the event was staged (see Stern, 1939). Method of recall (e.g., free recall versus structured questions) and length of time interval between the event and the recall attempt were often manipulated. The results of these studies were interpreted to mean that the accounts of eyewitnesses were neither very accurate nor complete. The research "eyewitnesses" were frequently quite accurate in their recall of the main details of the staged event, but they were often mistaken in the peripheral details and in estimations of height and weight. For reasons speculated elsewhere (e.g., Yuille and Wells, 1991), the *Aussage* researchers tended to emphasize the errors in their research subjects’ accounts. This trend is particularly evident in the practice of some of these researchers to consider omissions of details, no matter how peripheral or trivial, as errors in recall (e.g., Stern, 1910). On the basis of these studies, researchers also cautioned about the dangers of suggestive questioning, especially with children (e.g., Binet, 1900; Stern, 1910). A comprehensive overview of this period in the history of eyewitness testimony research can be found in Cutshall (1985).

Munsterberg added considerable impetus to the growing interest in the new "science of testimony" in North America with his move to Harvard University early in this century. Munsterberg was convinced that in just a few years *Aussage* research had produced sufficient knowledge to be of value to the criminal justice system. In his introduction to *On the Witness Stand* (1908), he noted that his "only purpose is to turn the attention of serious men to an absurdly neglected field which demands the full attention of the social community." (p. 12) Similarly, Stern (1910) argued that the new psychology would "provide a scientific basis for practical knowledge of, and judgments upon, human mental acts and qualities and must give assistance in the practical manipulation of human minds." (p. 270)

Responding to this call to apply their research findings, psychologists eagerly presented their conclusions to the courts. The first appearance of experts on testimony was in a murder trial in Belgium in 1911 (see Bartol and Bartol, 1987; Goodman, 1984). In this case, in which a serial killer had been preying on children, the only testimony implicating the accused was provided by some child witnesses. One of the expert psychologists for the defense, Varendonck, provided research findings to the court that emphasized the fallibility and malleability of children’s memory, opining:

"Everyone accused of crime based on children’s testimony does not encounter a psychologist to counter the indictment that lies heavily upon them....When are we going to give up, in all civilized nations, listening to children in courts of law?" (cited in Goodman, 1984, p. 27).

This expert evidence set the tone for the *Aussage* period and this general sentiment has continued into the contemporary era. The theme of
psychological testimony on eyewitness evidence, both then and now, is that eyewitness evidence should be viewed with suspicion. Another recurring issue, both then and now, was the equivocal manner in which these early psychologists presented their data. The early Aussage research, like contemporary research, was characterized by a lack of agreement in the researchers’ findings. For example, Munsterberg (1908) noted the following: “In some cases it was shown that the mistakes made after a week were hardly more frequent than those made after a day. Other experiments seemed to indicate that the number of mistakes steadily increases with the length of time which has elapsed.” (p. 54) However, despite the diversity of findings, psychologists such as Varendonck presented unequivocal conclusions in court.

Although psychologists such as Munsterberg, Stern, and Varendonck displayed great enthusiasm for the application of Aussage research, this enthusiasm was not echoed in the legal profession. In a caustic and insightful response, Wigmore (1909) attacked Munsterberg’s work on the grounds that the Aussage research base was deficient methodologically, the findings were often inconsistent, and jury-members were quite capable of evaluating testimony through common sense. Wigmore’s review of On the Witness Stand, which was presented in the Illinois Law Review in the form of a mock libel trial, examined the “new and exact methods” lauded by Munsterberg. Wigmore (1909) concluded that the methodology employed in the Aussage research was “still so little exact and so incapable of forensic use that even their well-wishers confess that 1000’s of experiments and years of research will be required before they will be practicable, if ever.” (p. 415) Considering the lack of consensus among researchers, noted above, as well as the sketchy nature of the research findings, Wigmore questioned the haste of Aussage psychologists to generalize their research findings to the courts: “The conclusions drawn from this and similar research seem to me in general to be over-hasty. Of course this research must be much more extended before the jurist can obtain any practical results for use in trials; for that purpose a long and systematic study will be required.” (p. 412).

Wigmore’s devastating review of Munsterberg’s work helped to force the Aussage research into a period of decline. As Hutchins and Slesinger (1929) observed, “Munsterberg was too confident and overshot his mark. The lawyer snapped back and put the psychologist in his place. The psychologist then retired from the courtroom and left the law to muddle on its own way.” (p. 14) The devastating effect of World War I on the discipline of psychology in Germany and the near total triumph of Behaviorism in American academic psychology also contributed to a virtual cessation of Aussage research (Bartol and Bartol, 1987).

As this short overview has demonstrated, this period of early activity by psychologists on the recall and recognition of witnesses was characterized by energy and enthusiasm within the discipline and skepticism in the legal profession. The lawyer charged with the task of evaluating the expertise of a proposed witness may find this brief history useful. Anyone who claims expert knowledge in the academic field of eyewitness memory should be aware of the Aussage period and the parallels between it and the contemporary scene.

THE REAPPEARANCE OF EYEWITNESS RESEARCH

Modern eyewitness research commenced in earnest in the early 1970s, primarily as a result of the work of two American psychologists, Elizabeth Loftus at the University of Washington and the late Robert Buckhout at the Center for Responsive Psychology in New York City. Close parallels to the Aussage era soon emerged in both the focus of the research—the fallibility and malleability of human memory—and the zeal of some psychologists to extend laboratory findings to the courtroom. Also, opposition to this extension surfaced in both legal and psychological circles. The hiatus of five decades had not altered the fundamental issue of debate: can the body of research findings on eyewitness testimony be generalized to the testimony of actual witnesses and victims of crime?

The impelling force for Loftus’ initial research was clearly theoretical. It appeared that her objective was to demonstrate the reconstructive nature of memory and considered the eyewitness domain to represent an ideal opportunity for such a demonstration. To accomplish this, Loftus created the “post-event misinformation” paradigm [although it was conceptualized earlier (e.g., Bartlett, 1932)]. In this paradigm (for a review see Loftus, 1979) the witnesses, usually undergraduate volunteers who are participating in the research for either course credit or money, are shown an event via slides or video tape. Following the event presentation, the “witnesses” are given a series of questions about specific details of the event. For the “witnesses” in the experimental group some of these questions include misleading information. Shortly thereafter both experimental witnesses and control witnesses (i.e., those who did not receive the misleading information) respond to a series of questions about the event. The critical concern is the extent to which the experimental witnesses incorporate the post-event misleading information into their account of the event. For example, many of these studies have involved the presentation of a slide sequence of an auto-pedestrian accident (Loftus et al., 1978). Following a short delay (usually seconds or minutes), the experimental witnesses were presented with a question containing misleading information, such as, “Did the car stop at the stop
sign?" when, in reality, the slides had depicted not a stop sign but a yield sign. On a subsequent recognition test, subjects were asked to choose between pairs of slides depicting a stop sign and yield sign, and they often inaccurately reported recognizing the slide showing the stop sign. In one study (e.g., Loftus et al., 1978) subjects who had been misled chose the correct slide 41% of the time compared to 74% by control subjects. The issues related to suggestibility of witnesses are discussed in greater detail in a later section of this chapter.

This post-event misinformation paradigm has become a favorite tool of researchers of eyewitness memory. The debate has raged for over twenty years about the nature and meaning of the effects found in this paradigm (e.g., McCloskey and Zaragosa, 1985). Scores of studies have indicated the following: the recall of experimental subjects can sometimes be changed by post-event misinformation if the focus of the information is peripheral rather than central in nature (e.g., Yuille, 1980), if the witness did not notice that aspect of the event at the time he or she witnessed it (e.g., Schooler and Loftus, 1986), and if the source of the misinformation is credible (e.g., Ceci et al., 1987). In short, this extensive research effort has demonstrated that witnesses may accept a reliable person's information about something trivial in an event if the witnesses didn't notice it at the time. An unresolved theoretical debate about these findings is whether the witness' memory is amended, i.e., changed by the misinformation, or appended, i.e., the misinformation is added to the memory (McCloskey and Zaragosa, 1985).

Continuing more directly with the theme of the Aussage research, Robert Buckhout's research was designed solely as a means of demonstrating the unreliability of eyewitness testimony. The focus of his research was on the application of findings to the criminal justice system and not on theoretical aspects of memory. His research also differed from that of Loftus in that his focus was on identification rather than verbal recall. For example, in one study Buckhout (1974) arranged for a local television news show in New York City to include a film of a staged crime in a newscast. A photo-spread was then displayed and viewers were encouraged to phone in their identification guesses. The chance performance that resulted from this task was taken as evidence for the fallibility of eyewitness identifications in actual criminal investigations.

The research of Loftus and Buckhout helped reawaken interest among psychologists in exploring eyewitness memory. In keeping with the Aussage era, soon after the re-emergence of eyewitness research psychologists began testifying in court, generally hired by the defense to discredit a witness in the eyes of the jury (Konecni and Ebbesen, 1986). A number of cases of this type are reviewed in Loftus and Ketcham's book, Witness for the Defense (1991). The courts have been open to hearing such testimony, although this varies as a function of jurisdiction. For example, although widely admitted in the United States, expert testimony on the competence of eyewitnesses is generally not admissible in Canada or the United Kingdom. The issues central to the debate concerning expert evidence are dealt with in more detail later in this chapter.

Since the re-emergence of eyewitness memory research in the 1970s, hundreds of research articles pertaining to eyewitness memory have been published in psychological journals. The following sections provide brief summaries of some of the different issues that have been addressed in eyewitness research in the past 25 years. For each issue, we summarize the central points, the weaknesses, or debates related to the research and provide the reader with suggested readings in that area of research.

INDIVIDUAL DIFFERENCES

Researchers have explored a variety of variables related to individual differences in eyewitness memory. These personological factors include gender, age, and intelligence. There is little consensus with respect to the effect of gender. Some studies report female superiority in eyewitness performance (e.g., see Shepherd, 1981 for a review); some report a male advantage (e.g., Clifford and Scott, 1978); and others report subtle male-female differences with no clear supremacy (e.g., Powers et al., 1978). This pattern of conflicting findings can give ammunition to critics who point to the inability of researchers in this field to develop methods that provide reliable results. However, research on gender differences across a range of topics (other than eyewitness memory) often produces conflicting findings (Hare-Mustin and Maracek, 1988). The general conclusion in much of this research is that greater variability exists within a gender (either males or females) than between genders. Thus, it is likely that the lack of conclusive findings related to gender in eyewitness research is a reflection of the difficulty of studying gender, which arises from a complex interaction of biological, psychological, and sociocultural factors.

There is no apparent systematic relationship between intelligence and eyewitness performance (e.g., Goetz, 1980). Even witnesses with intelligence levels substantially below average can, if carefully viewed, provide reliable accounts of events they have directly experienced (e.g., Bull and Cullen, 1992; Dent, 1986; Gudjonsson and Gunn, 1982). However, individuals whose intelligence levels are sufficiently low as to severely limit verbal abilities and communication skills may be compromised as witnesses. In summary, there is no clear evidence in the research literature that either gender or intelligence (with the exception noted above) are clear determinants of witnesses' reliability.
Unlike gender and intelligence, the age of the witness has shown a consistent relationship to eyewitness performance. When carefully interviewed children can be just as accurate as adults in recall (e.g., Goodman and Hahn, 1987; see Yuille et al., 1993). However, children are more susceptible to errors in identification tasks; they report less information than do adults (at least up to about 10 or 11 years of age); and they are more susceptible to the influence of leading questions (for a review see Ceci and Bruck, 1993a). As well, different populations of children may show different levels of suggestibility (e.g., deaf children; see Porter et al., 1995). There is also some research to indicate that elderly adults may be less accurate than other adults (e.g., Yarmey et al., 1983; see Yarmey, 1984 for a review). However, there are a variety of problems with the extant research on geriatric eyewitnesses. For example, only slide presentations have been employed; recall ability has been measured by written questionnaires rather than in an interview format; and subjects have been selected according to narrowly defined criteria. Any generalization of these findings to actual eyewitness situations is therefore suspect.

THE EFFECTS OF DELAY

The effect of delay on recall is one of the most controversial aspects of eyewitness research and one that should often be challenged in court. The controversy here arises from issues related to the context in which most eyewitness research is conducted and the fact that memory takes a different course over time, depending on the nature of the experience.

To clarify these concerns, the issues germane to generalizability of research findings are discussed in some detail below. As noted earlier, research is typically conducted in a university-based laboratory. Most psychologists prefer this context for several reasons:

1. It is convenient. It is relatively easy to obtain the cooperation of undergraduates to play the role of witnesses. Finding voluntary “witnesses” is obviously much more difficult in other contexts.

2. Psychologists are often interested in establishing cause and effect. The complexity of the “real world” makes this difficult. The laboratory provides simplicity by allowing greater control of the conditions of the experiment. As Banaji and Crowder (1989) observed, “our view is that the more complex a phenomenon, the greater the need to study it under controlled conditions, and the less it ought to be studied in its natural complexity.” (p. 1192)

3. In the laboratory the researcher knows with certainty what actually occurred, i.e., “ground truth” is established. In studies with real victims or witnesses, this is rarely the case.

The dependence on the laboratory or other controlled contexts requires the assumption that memory operates in the same fashion, regardless of context. The core of the debate, thus, revolves around the question of whether memory is context-free or context-dependent. Our reading of the literature is that memory is essentially a context-dependent phenomenon. That is, memory for an event will be affected by contextual factors such as the physical environment (e.g., Godden and Baddeley, 1975), one’s psychological state or mood at time of encoding and retrieval (e.g., Bower, 1981), personal relevance of the event (e.g., Hosch and Cooper, 1982), motivation to remember or forget the event (e.g., Tollestrup et al., 1994), degree of personal participation in the event (e.g., Yuille et al., 1994), and the effects of drugs or alcohol (see Read et al., 1992; Yuille and Tollestrup, 1990; Yuille, Tollestrup, Porter, and Marxsen, unpublished).

Psychologists have proposed that there are a variety of types of memory with each type serving a different function. In addition, different types of memory often exhibit a different forgetting curve, i.e., the passage of time affects each memory type in a different fashion. The most enduring division of types of memory is that developed by Tulving (1972, 1983), which includes procedural, semantic, and episodic memory. A more recent distinction has been made between implicit and explicit memory (e.g., Graf and Schacter, 1985), but given that this distinction has little relevance to the types of memories (i.e., episodic memories) of primary interest here, this distinction is not explored further. Procedural memory enables us to remember sensory-motor coordinations such as tying our shoes or riding a bicycle. We retain our definitional and functional knowledge about things in our environment (e.g., “a dog is an animal”) in our very enduring semantic memory. Many of our procedural and semantic memories remain relatively intact across the passage of time. A third type of memory, our ability to remember specific, personally experienced past events, is mediated by our episodic memory, the type of memory of greatest relevance to the criminal justice system and the one usually studied by eyewitness researchers. Within the domain of episodic memory, different memories are affected in different ways by the passage of time. The vast majority of episodic experiences, the routine events of life, are quickly forgotten. Because they are ordinary events, they quickly fade from memory, a process often called normal forgetting. However, certain occurrences, sometimes labeled “remarkable” events (Cutshall and Yuille, 1989; Yuille and Tollestrup, 1992), are not so readily forgotten. These are typically events that are unusual and may have had some special impact. These memories may be retained vividly, and perhaps quite accurately, for months or years, in fact, for a lifetime and may be “remarked” or reflected upon often. Thus,
normal forgetting may occur concerning many of the day-to-day events of one’s life, but a person may vividly remember a wedding, a first sexual experience, or a physical attack.

For obvious ethical reasons the events used by psychologists in the laboratory are relatively trivial. The research can have no negative or lasting effect on the witnesses. Thus, normal forgetting will typically operate for the experiences of the laboratory witnesses. Based on such research the psychologist may tell the court that witnesses quickly forget the details of a witnessed event. However, if a victim experiences a traumatic, "remarkable" event such as a sexual assault, normal forgetting may not apply. This type of episode may be vividly remembered. In such a case the psychologist, if he or she has only conducted laboratory based research, may have little or nothing to say that will be of assistance to the trier of fact.

The relationship between the witness studied in the lab and the witness who appears in court is a central issue in evaluating expert testimony by psychologists. The vast majority of researchers have never studied a victim or witness of an actual crime. They have never studied anyone who has been profoundly affected by a personally experienced event (the closest that some psychologists have come to this phenomenon is to study how people remember vicariously experienced events, like the assassination of John F. Kennedy or the Challenger disaster, e.g. Bohannon, 1988; Pillemer, 1984). The laboratory witness is an unaffected bystander to a personally innocuous event. There are no consequences of making a misidentification or providing inaccurate recall about the event (except to produce data subject to misuse by some psychologists!). Some witnesses in court also may be unaffected bystanders. For example, the victim of a credit card fraud may experience the event as just another routine sales transaction (Tollesrup et al., 1994) and thus may behave exactly like the laboratory witness. However, many of the witnesses who appear in court are victims who have been deeply affected by an event (Yuille, 1986). Because the consequences of his/her testimony may be profound, this witness may behave very differently from the laboratory witness.

The expert who is drawing conclusions about witnesses in one context (e.g., sexual assault, armed robbery, or physical assault) based upon research with unaffected bystanders should be rigorously challenged concerning the validity of generalizing between these two very disparate contexts. This point was concisely stated by Konecni and Ebbesen (1986): "The external-validity problems of the memory and perception research on which expert psychological testimony on eyewitness issues is based are so glaring that this type of testimony, at the present time, does not pass the Frye test." (p. 121)

There are a few psychologists that have studied the memory of witnesses to traumatic events (see Christianson and Nilsson, 1989; Cutshall and Yuille, 1989; Fisher et al., 1989; Sporer, in press; Tollesrup et al., 1994; Yuille and Cutshall, 1986, 1989; Yuille and Kim, 1987—these studies are reviewed in some detail below). Experts who are aware of this work and are able to differentiate the applicability of research findings to different contexts should be of much more use to the courts.

SUGGESTIBILITY AND MISINFORMATION

Some years ago, Wells (1978) suggested an important dichotomy about the kinds of variables psychologists study in eyewitness research. Variables that may have an influence on a witness at the time of the event (for example, lighting conditions, distance from the event, duration of the view of the perpetrator, the presence of a disguise) were labeled as "estimator" variables. That is, psychological research could estimate the average effects of such variables but could not alter their impact. Alternatively, variables that affect witnesses during the investigation (e.g., the nature of a lineup, the instructions given to a witness before an attempted identification, the form of the interview with the witness) were labeled "system" variables. These aspects of the system could be modified to maximize the likelihood of reliable eyewitness performance. Suggestibility and misinformation are influenced by system variables.

As noted earlier, the research of Loftus and her colleagues introduced the post-event misinformation paradigm. There has been a great deal of debate about the results from this paradigm. Some, particularly Loftus (e.g., 1983), have argued that the laboratory demonstrations reflect a fundamental property of memory—it is highly malleable. Others (Bregman and McAllister, 1982; Dodd and Bradshaw, 1980) have suggested that the dynamics of a real-life eyewitness situation greatly reduce susceptibility to misleading information and suggestion. Indeed, Yuille and Cutshall (1986) found that they were unable to mislead actual eyewitnesses regarding the peripheral details of a shooting they witnessed.

As a system variable, suggestion can be addressed best by assuring that victims and witnesses are interviewed in a non-leading, non-suggestive fashion (Yuille et al., 1993). However, if interview or investigative procedures become a focus of expert evidence on the effects of suggestion, the expert should acknowledge the diverse findings concerning suggestibility effects. Further, the expert should emphasize that the effects of leading questions may vary with context.

REPRESSION AND THE "FALSE MEMORY" DEBATE

One form of suggestion has received a great deal of attention in the last few years, in the context of allegations of childhood sexual abuse by
adults. As awareness of the widespread sexual exploitation of children has grown (e.g., Finkelhor, 1984), a literature has developed that is based upon the assumption that many adults who were sexually abused as children, particularly where the abuse involved incest, have "repressed" their memory for the abuse (e.g., Bass and Davis, 1988). There has been a strong reaction against this assertion including a group, the False Memory Syndrome Foundation, claiming that therapists are creating false memories of childhood abuse in their clients. The claim is that through the use of suggestive techniques therapists are causing the development of false memories in vulnerable patients.

Repression is a theoretical term related to amnesia. Amnesia represents a form of relatively rare memory loss. Unfortunately, the term amnesia is often used loosely or informally as if it were a normal aspect of human memory rather than a pathological condition (Rubinsky and Brandt, 1986). For example, in a case in Arizona, amnesia was confused with everyday forgetting when a U.S. court ruled that since everyone has experienced episodes of forgetfulness at one time or another, "everyone is amnesic to some degree" (State vs. McClenon, 1968). This represents a gross misuse of the neuropsychological term, amnesia, and it reflects an ignorance of the syndromal aspects of pathologically impaired memory demonstrated in decades of experimental and clinical research (Wiggins and Brandt, 1988).

Two classes of amnesias, organic and psychogenic, have been distinguished in the psychiatric and psychological literature. Organic amnesia represents a large class of amnesias associated with a wide variety of neurological conditions, including alcoholic Korsakoff syndrome, Alzheimer's, Huntington's disease, head trauma, strokes, or brain tumors. Organic amnesiac individuals may suffer from anterograde amnesia, an inability to record ongoing events and learn new information, and/or from varying degrees of retrograde amnesia, difficulty in recalling events that occurred prior to the onset of the amnesic condition. The latter form of amnesia (retrograde amnesia) is more commonly seen, for example, in the motor vehicle accident victim with mild head injury who is unable to recall events immediately preceding the injury (usually only a few seconds are permanently lost). There are also forms of organically based temporary or transient amnesia; for example, amnesia induced by drug or alcohol intoxication or by epileptic seizures.

The more controversial class of amnesias, of particular interest here, are those known as dissociative or functional amnesias. As described by the American Psychiatric Association's Diagnostic and Statistical Manual (DSM-IV), this disorder is characterized by a sudden inability, with no organic basis, to recall important personal information. (The previous Manual, DSM-III-R, called this phenomenon psychogenic amnesia.) The extent of this memory loss is too great to be explained by ordinary forgetfulness. Of the five types of dissociative amnesia, the most common type, according to DSM-IV, is localized amnesia, where there is failure to recall all events occurring during a circumscribed period of time, usually the first few hours following a profoundly disturbing event. Somewhat less common is selective amnesia, the failure to recall some, but not all, of the events occurring during a circumscribed period of time. The other three forms are relatively uncommon: generalized amnesia encompasses the whole life and is often fodder for the Hollywood gristmill; continuous amnesia, as its name implies, continues into present; systematized amnesia refers to loss of memory for a category of information, such as memories related to a particular person. Dissociative amnesia begins suddenly, usually following severe emotional distress, and termination of the amnesia is typically abrupt (e.g., Schacter and Wang, 1982). Although Dissociative Amnesia is categorized in DSM-IV as a specific mental disorder, episodes of dissociative amnesia are often associated with Post-traumatic Stress Disorder and Dissociative Identity Disorder, two disorders often diagnosed for individuals with a history of child sexual abuse (Herman, Perry, and van der Kolk, 1989; Landecker, 1992; Peterson, 1991).

The concept of repression, which has become a focus of interest in the current debate about the veracity of adults' recollections of childhood sexual abuse, is related to, but not synonymous with dissociative amnesia. Repression is a psychoanalytically-based theoretical concept that was introduced by Freud and has been postulated as a mechanism to explain dissociative amnesia, especially localized or selective amnesia following an extremely traumatic experience such as sexual abuse. Theoretically, repression represents an unconscious process by which the human psyche protects or defends itself against overwhelming fear, loss, or anxiety by removing or hiding an intolerably painful and/or frightening experience from a person's own conscious awareness. An incident is, thus, "forgotten." Although this process may seem somewhat akin to the motivated forgetting mentioned previously, it differs in some very critical ways. Repression is hypothesized to be an unconsciously motivated process (i.e., outside of an individual's awareness) and the unconscious, repressed material (e.g., sexual abuse) supposedly affects the ongoing functioning of the individual (e.g., psychological symptoms such as depression, eating disorders, low self-esteem). Motivated forgetting, on the other hand, is a deliberate process of which the individual is consciously aware; and, no assumptions are made concerning the relationship between the forgotten material and current symptomology.

There is no firm empirical support for the theoretical construct of repression (e.g., Holmes, 1990), perhaps because such a concept is not
amenable to experimental verification or refutation. It is therefore important to distinguish between dissociative amnesia, an acknowledged type of memory loss, and repression, an unverifiable, hypothetical explanation for dissociative amnesia. Amnesia has not been well researched; "dissociative disorders (including dissociative amnesia) remain among the most poorly understood clinical syndromes" (Davison and Neale, 1986, p. 162). What is known is that psychogenic or dissociative amnesia exists, but there is little agreement beyond this fact. There is no firm knowledge concerning the incidence rates, the circumstances in which amnesia is likely or unlikely to occur, or the type of person who might be prone to amnesia.

A similar lack of knowledge exists concerning the phenomenon of created memories. There is evidence that it is possible to create a false memory of something that never happened (e.g., Loftus, 1993). However, beyond this fact there is no knowledge about the incidence rates of false memories (i.e., how common or uncommon they are as a result of therapy), if certain clients are more susceptible than others, if created memories are qualitatively different from "real" memories, or if the pattern of recovery of valid memories is different from the pattern of creation of false memories. There has been a recent attempt to delineate those therapeutic practices that may increase the risk of creating false memories (Lindsay and Read, 1994).

Expert evidence in this area must be very carefully evaluated. There is considerable polarization in this debate between those who believe that repression is widespread and those who believe that repression does not exist. An "expert" who belongs to one or the other of these opposing groups may provide a distorted or exaggerated impression of the state of knowledge in this field. This is an area in which some psychologists and other mental health professionals seem willing to go well beyond the current level of knowledge and draw broad, unwarranted conclusions.

THE EFFECTS OF STRESS

No variable captures the debate concerning the applicability of eyewitness research to the criminal justice context more strongly than stress. Research into the effects of stress has yielded a confusing pattern of results. For example, Deffenbacher (1983) reviewed 21 studies and noted that ten of them suggested that stress either increases eyewitness accuracy or has no effect, whereas the findings from eleven studies led to the conclusion that stress decreases eyewitness accuracy. It is interesting to note that of the former ten studies, the majority involved staged live events, whereas of the latter eleven studies suggesting that arousal decreases accuracy, only one study employed a live event.

Part of the problem here is the definition of "stress." In some studies, arousal of the witnesses was induced with the application of white noise (a loud hiss presented via earphones) or an electric shock. In these studies, the source of the stress is completely independent of the event to be remembered. One of the things we have learned about stress is that it tends to focus attention (e.g., Easterbrook, 1959; Eysenck, 1982; Loftus, 1980; Mandler, 1975). That is, as someone becomes stressed, his/her attention tends to narrow to either the cause of the stress and/or the effects of the stress. Thus, studies that use a stressor that is external to the eyewitness event are actually studies of unpleasant distractors rather than effects of witnessing or experiencing a stressful criminal act.

Other studies have employed films as stimuli (e.g., Clifford and Scott, 1978; Clifford and Hollin, 1981; Loftus and Burns, 1982). In this type of research a witness’s stress is created by the content of the film. Thus, some subjects may be shown a film with a gun shooting at the end of the film (stressful condition), whereas others view the film without the shooting (control condition). It is astonishingly naive of psychologists to believe that most undergraduates would find the content of such films stressful. Most people see so many shootings and other extreme violence in movies and on television that they are inured to it (Myers, 1990, reported that by high school graduation the average American has witnessed about 25,000 violent deaths on television and in films). Studies that employ films as stimuli generally bear no relationship to the impact of real life stress on victims of crime. Any testimony based on such research that claims applicability to victims or witnesses of actual crimes should be dismissed on common sense grounds. As Christianson (1992b) notes in a recent review, "it may be that inconsistencies in empirical findings reflect differences in the to-be-remembered (TBR) detail information and testing circumstances." (p. 308)

Some psychologists invoke the Yerkes-Dodson law to explain the effects of arousal on eyewitness memory (Mandler, 1992; see Christianson, 1992a). This "law" was proposed early in this century to describe the relationship between arousal and performance in rats in simple learning tasks (Yerkes and Dodson, 1908). The relationship, an inverted U-shape, summarizes the fact that animals' performance improves on learning tasks as arousal increases until an optimum level of arousal is reached. After this point, increases in arousal have a detrimental effect on performance. The application of this principle to eyewitness performance is unwarranted and without empirical foundation (Christianson, 1992b).

Recent studies of those who have experienced high levels of stress (e.g., victims of armed robbery; see Tolleserup et al., 1994) indicates that the impact of stress is very complex. Some victims of stressful events may have vivid and detailed memories of the events, whereas others may
only remember their subjective feelings of stress or may even suffer amnesia for the events (Yuille and Tollesstrup, 1992). The complexities of opinion on this issue can be appreciated by examining Christianson’s (1992a) text, *The Handbook of Emotion and Memory*. “General statements (concerning the relation between memory and stress) seem unwarranted both in the literature and in practical settings” (Christianson, 1992b).

**EYEWITNESS IDENTIFICATION**

The ability of witnesses to identify someone from a lineup or photo-spread has been the focus of a considerable amount of research, as well as an issue that frequently brings psychologists to court. In a sense, the problems related to identification are similar to those raised in the section on suggestibility. Indeed, a lineup or photo-spread may be considered a leading question: the presentation may suggest to the witness that he/she is expected to make a choice. If the culprit is not present in the lineup and an innocent suspect is selected, a serious miscarriage of justice could ensue.

Several recent articles have provided thorough reviews of this area of research (Wells, 1993; Wells and Turtle, 1987; Tollesstrup et al., 1994). Many estimator variables have been identified that affect the ability of witnesses to correctly identify the culprit: the lighting conditions under which the culprit was originally viewed, the length of time of originally viewing the culprit, the distance of the witness from the culprit, the focus of the witness’ attention, the length of time between the original event and the attempted identification. The race of both the witness and the culprit may also affect the identification (see Brigham and Malpass, 1985 for a review), although some contention remains about this issue (Lindsay and Wells, 1983). The identification may also be influenced by events between the crime and the attempted identification. For example, witness accuracy may be diminished if the witness previously viewed a set of mugshots or if the suspect has been viewed in some other context (e.g., Loftus, 1976; Ross et al., 1994).

It is important to note that recognition memory operates differently than recall. For the identification task, the witness needs only to believe that he or she has seen the suspect before. This recognition response may occur because the suspect was the culprit, the suspect was seen by the witness in some other context than the crime, or the suspect shares enough similar features with the culprit to trigger a false recognition response. A number of cases of false identification have been documented (e.g., Buckhout, 1974; Devlin, 1976; Loftus, 1979, 1986).

Expert testimony concerning these estimator variables may be useful to the triers of fact on some occasions. In contrast to recall memory, there is little reason to suspect that the perceptual process of recognition occurs differently in the criminal context than in the laboratory (although the untested possibility exists that the forensic witness might focus longer on each member of the lineup to ensure a correct decision; see below). However, because the consequences of a recognition choice in laboratory studies is trivial, a fundamental difference in decision-making exists. The selection of a choice from a photo-spread or lineup is rather like a game in the laboratory, whereas it is a very serious act in the criminal context. The contrast in the consequences of identification in the two contexts could improve performance in the criminal situation; real-life witnesses could be more concerned about the consequences of making a false identification in a criminal matter. Alternatively, recognition errors might increase with victims of actual crimes who have a strong desire to have someone punished for the crime. Unfortunately there has been little research on this issue (see Malpass and Devine, 1981).

Recent research on identification has focused more on system variables, especially on the form of presentation of photo-spreads. First, psychologists have been trying to develop methods for evaluating the fairness of a photo-spread or lineup (see Wells, 1993). The most common method currently in use is to provide nonwitnesses with a simple description of the culprit and then ask if they can identify him or her in the photo-spread or photo of the lineup. The assumption is that if the lineup was fair, all the participants in the lineup should be picked with roughly equal frequency by nonwitnesses using only a verbal description of the culprit as an aid. Wells et al. (1979) define the “functional size” of a lineup as the number of participants fitting the description of the target. This is compared to the “nominal size,” which is simply the number of warm bodies in the lineup. In short, the suspect and the foils in the photo-spread or lineup should all roughly fit the same description; the functional size should equal the nominal size of the lineup. Although this technique has encountered some criticism in the field, it is widely used to assess lineups in forensic contexts. However, if all of the lineup’s members bear a very great resemblance to each other, the lineup is then said to have lost its “propitious heterogeneity” (Luus and Wells, 1991). In such cases, identifying the target among the highly similar distractors could be a very difficult task.

Psychologists have also been concerned with trying to reduce false identifications of innocent suspects. They have suggested that the following methods can aid this goal (see Wells and Turtle, 1987; Wells, 1993):

1. The witness must be instructed that the culprit may not be present in the lineup or photo-spread.
in a debate with Elizabeth Loftus (1983) that was published in the flagship journal of the American Psychological Association, *American Psychologist*. Repeating the lamentations of Wigmore (1909), McCloskey and Egeth (1983) argued that psychologists do not yet possess a body of valid, replicable research that goes beyond the knowledge of the juror. They further suggested that the laboratory base of most psychological research in this field compromised any ability of psychologists to testify in court. Loftus responded by arguing that jurors do not possess the requisite knowledge of how particular psychological factors influence perception and recollection and thus they need this information from psychologists. Unfortunately, considering the base of our knowledge, this argument is little more than an assertion that the blind should be led by the blind. In an even more damning article, Faust and Ziskin (1988) argued that expert testimony by psychologists in general is questionable given the unreliability of clinical diagnosis and research. Furthermore, caution in presenting expert testimony is called for given the potential impact such testimony may have. It has been well-documented that expert testimony on witnesses can have a significant impact on juror decision-making (Bank and Poythress, 1982; Cutler et al., 1989; Fox and Walters, 1986; Hosch et al., 1980; Loftus, 1980; Saks, 1990; Wells, 1986; Wells et al., 1980; Yuille, 1989). Nonetheless, psychologists have continued to testify as experts regarding eyewitness testimony, particularly in the United States. Observing this ongoing trend, many critics have been advocating caution and the exigency for further research in naturalistic contexts (Yuille, 1989; Yuille and Wells, 1991, Yuille, 1993). Echoed by Elliot (1993), Yuille (1989) maintained that “we need to discourage strongly the court presentation of premature and unsupported ideas.” (p. 194)

Some have simply argued that the laboratory is preferred because it permits controlled research. “The implication that tests in the real world permit greater generalizability is false once the immense variability from one real world situation to another is recognized” (Banaji and Crowder, 1989, p. 1189). The reality of the complexity of real world dynamics is dismissed as a problem and used as an excuse to retreat to the comfort of the laboratory. Others have asserted that it is in the hands of the critics to demonstrate the lack of validity of the research before attacking the expert witness on that basis: “we do not have the luxury of waiting until researchers get around to completing all the studies that would be desirable” (Loftus, 1986, p. 249). This latter argument is specious and indefensible. No one is hurrying the psychologist into the courtroom except the psychologist. As Konecni and Ebbesen (1986) observed:

How do the advocates of testimony know that our discipline and our fund of findings are “ready” now to enter the courtroom? It would seem that their
claims may have less to do with the actual, proven maturity of psychology and its findings than with these researchers-testifiers' personal and professional maturity, stature in their field, and accomplishments: They are ready, so the results better be ready as well. (p. 122)

The most appropriate stance seems to be that lack of ecological validity must be assumed. Each issue for which the expert is tendered to the court must be addressed on its own merits. Does the psychologist have meaningful research from an appropriate context on which to base valid opinions for the court? If the answer to this question is positive, the expert may be of assistance to the court.

RESEARCH OUTSIDE THE LABORATORY

Perhaps psychologists find themselves in the strongest position to generalize their findings when they have researched a particular problem in a variety of contexts. Noting the near exclusive adherence to the experimental laboratory study by eyewitness researchers, Davies (1990) stated, "no one research method can of itself provide a reliable data base for legislation or advocacy. Rather, problems need to be addressed from a number of perspectives, each of which makes a different compromise between ecological validity and methodological rigor." (p. 21) Other approaches to researching eyewitness testimony outside of the laboratory include the field study, the case study, and archival research. Field studies are generally quasi-experimental studies taken out of the context of the laboratory. The subjects are usually not the "bread and butter" of laboratory experiments—the university undergraduate—but are drawn from other populations. These subjects are frequently unaware that they are participating in an eyewitness study. Often, random assignment to an experimental condition, a requirement for any true experiment, can not be attained. Case studies look at the memory of actual eyewitnesses to a real crime, typically from a single case. Often the police files are examined and the eyewitnesses reinterviewed for the study. Archival research examines police records concerning eyewitness accounts of a number of crimes. Unfortunately, all of these methods have been practically ignored by researchers in the area (Tollestrup et al., 1994). Even more worrisome, the few studies conducted with these methodologies have yielded many results that contradict consistent experimental laboratory findings.

1. Field Studies. The field study has become a popular alternative to the laboratory during the past decade. Psychologists have used police trainees (e.g., Yuille et al., 1994), clerks in malls (e.g., Read et al., 1990), and convenience store clerks (e.g., Brigham et al., 1982; Kafka and Penrod, 1985) in attempts to conduct field studies of witnesses. For example, in a study exploiting the unique conditions of the London Metropolitan Police Training Centre in Hendon, England, Yuille et al. (1994) examined the effects of stress and involvement on eyewitness recall. The Metropolitan Police Training Centre contains a mock building that the trainers "populate" with actors. Recruits must go into this very realistic simulation and perform police duties while interacting with the actors. In interviews twelve weeks after the events, police recruits who were asked to recall a stressful event (an unpleasant confrontation) were actually more accurate than those who were asked to recall a non-stressful event (a rather mundane interaction). In regard to the amount of information given, however, recruits recalling the non-stressful event gave more information than those relating the stressful event, despite the fact that the stressful event lasted longer than the non-stressful event. Those recruits who actually participated in the event, regardless of it being stressful or non stressful, recalled more after twelve weeks than those who merely observed the event. The effects of stress, delay, and participation were all different from expectations founded on laboratory research. Fisher et al. (1989) conducted a field study (the main purpose of which was the examination of the effects of training police investigators in interviewing techniques) and found that (regardless of whether the interviewer had received training) there was a surprising 94% agreement of those eyewitness statements that could be corroborated by another eyewitness—very unlike the accuracy rates typically found in laboratory studies.

2. Case Studies. In one of the earliest studies to utilize an alternative methodology, Yuille and Cutshall (1986) examined accounts of eyewitnesses to a shooting incident during the robbery of a gun store. This case study consisted of both an examination of Royal Canadian Mounted Police interviews of the 21 eyewitnesses (including the owner of the shop who, after being shot by the robber, returned fire and killed the thief) and interviews conducted by the researchers with 13 eyewitnesses who consented to participate. These research interviews were conducted between four and five months after the crime. The witnesses were very accurate in their original accounts and there was little change in either amount or accuracy after the delay of several months. The errors that were found consisted of recall of colors and estimations of numerical details (e.g., height, weight, age or number of gunshots) and errors accountable by the witnesses' vantage point. Stress (conceptualized here in terms of proximity to the gun fire) seemed to have no negative effect on the eyewitnesses' accounts.

3. Archival Studies. Archival studies provide another example of the contrast between laboratory results and those from other contexts, for example with the phenomenon of weapon focus. Weapon focus is the
CREDIBILITY ASSESSMENT

Reflecting the focus of the field of eyewitness research over the past century, this chapter has focused upon the debates concerning the accuracy and detail of eyewitness memory. The concern of the research has always been with the malleability and fallibility of witnesses who are trying to tell the truth. Very little attention has been devoted to techniques to distinguish truthful from fictitious accounts. However, recently North American researchers have become aware of a European practice of assessing eyewitness credibility that has been in use for over four decades (for histories of the development of these procedures see Trankell, 1972; Weger, 1989; Joffe, 1992). The procedures, variously called Statement Reality Analysis, Statement Validity Analysis or simply Statement Analysis, are based upon the assumption that the descriptions of actual experiences are qualitatively and quantitatively different from the description of an invented episode. The procedure consists of two components: (1) content-analysis in which a statement is examined for the presence of features associated with actual experiences, and (2) examination of other evidence in the case. Nineteen criteria are used to evaluate the credibility of the statement (e.g., quantity of details, logical structure, accounts of subjective mental state; see Raskin and Yuille, 1989). For example, a high degree of detail is a good indication of credibility in that it is difficult to embellish a false testimony with details because they do not exist in memory and because of the increased difficulty of maintaining a story’s consistency. It is argued that someone trained in this approach can reliably determine whether a statement is based upon experience or invention (Yuille, 1988; Landry and Brigham, 1992; Porter and Yuille, 1995).

Recently, there have emerged three experimental examinations of the utility of statement analysis techniques with adult eyewitness or suspect accounts (Landry and Brigham, 1992; Porter and Yuille, under review; Zaparniuk et al., in press) and one theoretical examination (Porter and Yuille, 1995). Each of these studies concluded that verbal clues can be of some use in distinguishing truthful from invented accounts. However, this represents a very modest research foundation and much more work must be done to assess the value of these techniques (Wells and Loftus, 1991). It is worth noting that statement analysis is routinely used in Germany to assess the veracity of uncorroborated allegations made by adults. Also, the courts in that country routinely accept expert evidence opinion about witness credibility based upon the use of statement analysis techniques.

CONCLUSIONS

It is obvious from this review that, depending on the research paradigm employed, widely varying conclusions can be drawn about specific eyewitness factors. It is unacceptable for experts in eyewitness testimony to offer opinions about the performance of forensic witnesses based solely upon laboratory based research. At the very least, experts must specifically outline the limitations and qualifications to the applicability of their research. Indeed, the recommendation of the Committee on Ethical Guidelines for Forensic Psychologists (1991) state that “Forensic psychologists (should) take reasonable steps to correct misuse or misrepresentation of their professional products, evidence, and testimony.” (p. 663) Thus, experts should be expected to justify empirically the appropriateness of their findings to the criminal justice system. The assumption that laboratory research findings apply to the forensic context should be challenged.

Although the cautions raised here are of importance to the fairness of the judicial process, a more fundamental issue is raised by the difficulties of presenting fair, objective expert evidence in an adversarial system. It is our position that the most balanced evidence is presented by the expert who appears as a “friend to the court” rather than as an advocate for one particular position (for an alternative viewpoint see Diamond, Slovenco; Chapter 1 in this edition). Given the complexity of most psychological data, an advocate of one position will likely bias his or her presentation of information in a way that obscures, rather than illuminates, the issue under consideration. In addition, an expert who attempts to present a balanced picture that does not go beyond the knowledge base of our field will enhance his or her credibility. Within an adversarial system, this may often be an elusive goal; however, it is a goal toward which psychologists as expert witnesses should strive.

Psychologists do have useful and valid information to assist the triers of fact. Memory is a complex process and many of the features of memory go well beyond common sense and experience. The courts will
benefit from expert testimony that fully represents the current state of knowledge in the field. The discipline of psychology will also benefit when the eagerness to apply knowledge is tempered by an appreciation for the contextual nature of human behavior and memory.

AUTHORS' NOTE
The authors would like to express appreciation to Melissa Nisbett who helped in the preparation of this manuscript.

REFERENCES


Coping with Psychiatric & Psychological Testimony

Coping with Psychiatric & Psychological Testimony

rent Trends and Developments (pp. 182-200). New York: Cambridge University Press.


**CITATIONS**
