

Chapter 9: Limitations of Expert Psychology Testimony on Eyewitness Identification

Heather D. Flowe, Kristin M. Finklea, and Ebbe B. Ebbesen

University of California, San Diego

Address ms. Correspondence to:

Heather D. Flowe
Department of Psychology, 0109
University of California, San Diego
9500 Gilman Drive
La Jolla, Can 92093

Chapter 9: Limitations of Expert Psychology Testimony on Eyewitness Identification

Psychologists are being called upon more than ever to evaluate the reliability of eyewitness testimony. This state of affairs is novel because the legal system's reception of psychological research was cold early on, stemming in part from the contention that the research was lacking in legal verisimilitude (Yarmey, 2001; also see Epstein, Chapter 4, this volume). The frustration felt by psychologists regarding the initial unwillingness on the part of legal officials to apply psychological research findings in the courtroom and in the police station is well-exemplified by Wells, Memon, and Penrod (2006), who write, "psychologists were able to use experiments to identify eyewitness problems long before the legal system was smacked in the face with DNA exonerations" (p. 46).

The external validity of eyewitness testimony research has also been debated within the field of psychology (see the 1986 special issue of *Law and Human Behavior*, McCloskey, Egeth, & McKenna). Some have argued that the conditions (the witnessing, testing, and/or measurement contexts) experienced by laboratory research participants may not be representative of the conditions faced by actual witnesses (e.g., Clifford and Lloyd-Bostock, 1983; Ebbesen & Konečni, 1997; Elliott, 1993; Konečni & Ebbesen, 1986; McCloskey & Egeth, 1983; Yuille, 1993), and others have argued that the methods and subject samples employed by researchers are diverse enough (e.g., Deffenbacher, 1984; Haber & Haber, 2000; Loftus, 1983; Yarmey, 1997; Yarmey, 2001) and the body of research evidence extensive enough (see Pezdek, Chapter 2, this volume) to warrant the generality of the findings. We will revisit some of these issues in this chapter, and argue in the main that the methods by which research findings are translated to legal cases by expert consultants raise important issues that require careful consideration.

In the sections that follow, we will argue that generalizing from laboratory research to actual cases we must consider 1) the extent to which the procedures or psychological processes that instantiate variables in the laboratory occur outside of the laboratory; 2) whether the background conditions of laboratory studies are diverse enough to warrant gross application; 3) whether testifying about a given factor (e.g., weapon presence, cross-race witnesses, post identification confidence) provides incremental validity over traditional safeguards; and 4) the extent to which the process by which the legal system selects witnesses to testify affects the generalizability of laboratory research to a criminal case. Experts also have been known to argue to the court that the DNA exoneration cases provide evidence that eyewitness memory is unreliable (e.g., *People v. Adams*, 2008; *People v. Copeland*, 2007; *People v. Davis*, 2004; *People v. Ellison*, 2008; *People v. Garcia*, 2008; *United States v. Burton*, 1998). In this chapter, we will comment on the limited generalizability of the DNA exoneration cases (for other examples of criminal cases in which DNA exonerations were discussed in relation to eyewitness identification reliability see: *People v. Dubose*, 2005; *People v. Herrera*, 2006; *People v. Lawrence*, 2007; *People v. Lewis*, 2008; *People v. Romero*, 2007; *People v. Williams*, 2006; *People v. Woolcock*, 2005). Finally, we will argue that by acting as case consultants in an adversarial role, we possibly hamper our future ability to work with the legal system as scientists in developing procedural safeguards to improve the reliability of eyewitness testimony.

At the outset, it is important to point out that expert testimony on eyewitness identification can influence the disposition of criminal cases in a number of ways. Typically one thinks of an eyewitness expert as someone who educates the jury with respect to the factors that can influence eyewitness memory. Expert opinion, however, can influence the disposition of a criminal case at other stages. In our interactions with the legal system, we have learned that expert opinion can be

influential in plea bargaining a case, in deciding whether to admit the testimony of a particular witness, and in preparing a case for appeal. (Unfortunately, there are no systematic data on these issues.) In the sections that follow, our discussion applies to not only expert testimony proffered to juries, but also to eyewitness expert opinions offered in specific cases at other stages processing.

The procedures and/or psychological processes that instantiate a given variable in the laboratory might not arise outside of the laboratory

How do expert witnesses on eyewitness memory generalize from laboratory research to a real world case? They do so by comparing the similarity of attributes found in a specific case to the conditions simulated in laboratory studies:

“To the extent that a particular variable or set of variables was present at the time of the witnessing, and those variables are known through scientific research to have a particular effect or set of effects, then it is possible that expert testimony on those witnessing conditions could be relevant and of assistance to the jury in evaluating particular points at issue” (Yarmey, 2001, p. 94).

The above statement makes the assumption that research variables themselves are causal entities, rather than components that are tied to specific laboratory procedures. Danziger and Dzinis (1997) traced the historical use of the term “variable” in psychological research by empirically analyzing the context in which it appeared in psychology papers. They found that the term was first imported into psychological research in the 1920’s to describe statistical procedures. Subsequently, variables began appearing in papers as methodological entities to describe the procedural aspects of experiments. Variables eventually graduated into theoretical and substantive entities in discussions of research findings. In explaining this transformation, they propose, “That may point to a fairly widespread, though implicit and unexamined, belief that, any psychologically relevant part of reality was already prestructured in the form of distinct

variables, and that psychological research techniques merely held up a mirror to this structure” (Danziger & Dzinis, 1997, p. 46).

If the procedures that give rise to a construct in the laboratory differ from the procedures that give rise to a construct in the real world, then simply matching variables found in the scientific literature to the variables found in a criminal case based on prima facie resemblance may be inappropriate. To illustrate, *stress* as a construct may manifest differently in the real world than it does in the laboratory. Stress may be heightened by perceived danger, which could vary depending on the proximity of the eyewitness to the perpetrator in the real world (Civilini & Flowe, 2008; Yuille & Cutshall, 1986). It is plausible that being in closer proximity to the perpetrator may offset any negative impact of stress on a given witness’ memory. Additionally, the onset of stress in relation to when the perpetrator is first seen by the eyewitness may also be important, and research has not addressed this issue. More than a third of witnesses in our field research saw the perpetrator for a period of time prior to being victimized (Civilini & Flowe, 2008). Additionally, the procedure that is used to instantiate eyewitness recall in the laboratory may also be an important determinant of whether stress has a negative effect on memory reports. Deffenbacher et al. (2004) reviewed the laboratory literature on the impact of stress on the accuracy of eyewitness memory. Their meta-analysis indicated a statistically significant increase in errors in the stress compared to the control conditions for descriptions that were given under *interrogative recall*. The error rates in the stress and control conditions did not differ under conditions of *free recall*. These results indicate that the effect of stress on description accuracy may be tied to other psychological variables (e.g., interference, fatigue) that are affected by the type of interview procedure that is used. Though the procedure police use to interview witnesses may vary across jurisdictions, field research suggests that witnesses typically provide testimony

in conditions that are more similar to free recall than interrogative recall (Civilini & Flowe, 2008; Fahsing, Ask, & Granhag, 2004; Lindsay, Martin, Webber, 1994; Sporer, 1996). The extent to which the effects of stress reported by laboratory studies occur in real world cases may depend on the extent to which the interview conditions found in real world cases matches those of the laboratory. The Deffenbacher et al. (2004) meta-analysis also indicated that in the stress compared to the control conditions there was an increase in filler identifications in *target present* lineups. No effects were found in *target absent* lineups. Wells, Memon, and Penrod (2006) discuss the importance of including perpetrator absent lineups in laboratory investigations, as “target-absent lineups simulate the real-world situation in which police have focused their suspicion on an innocent suspect” (p. 50). The variables that control the error rate in target present and target absent lineups may differ; therefore, testimony about factors that lead to errors in target present lineups might be completely irrelevant in cases in which the issue is whether the witness has mistakenly identified an innocent suspect.

As another example, research has found that misleading post event information has a negative impact on memory reports (Ayers & Reder, 1998). An expert may be inclined to draw broadly from this research literature and testify that eyewitnesses who encounter misleading information are more likely to have inaccurate memory reports. Laboratory research has indicated, however, that the misinformation effect is affected by the procedures that are used to introduce the misleading post event information, including: the timing of the presentation of the misinformation (e.g., Loftus, Miller, & Burns, 1978), the retention interval between misinformation presentation and the final memory test (e.g., Loftus, Miller, & Burns, 1978; Windschitl, 1996), whether the misleading information is about central versus peripheral details (e.g., Wright & Stroud, 1996), the format of the questions on the final memory test (e.g.,

McCloskey & Zaragoza, 1985), and the type of source providing the misinformation (e.g., Hope et al., 2008). Whether these effects manifest in a real world case depends on the totality of circumstances involved, and broad generalizations (e.g., testifying that “research generally finds that exposure to misleading information can reduce the accuracy of memory reports”) may be misleading.

A related issue is whether psychological construals of a given laboratory procedure can be broadly applied to the legal system (see Berkowitz & Donnerstein, 1982 for a discussion of psychological construal as it relates to the external validity of social psychology experiments). To illustrate, the *own-race bias* describes the finding that the rate of making a correct identification is lower if the to-be-remembered face is of a different race than the participant eyewitness (e.g. Brigham & Malpass, 1985; Feingold, 1914; Meissner & Brigham, 2001; Sporer, 2001). The name of the effect alludes to the causal mechanism being tied to a difference in race between the witness and perpetrator. Currently, the most widely supported explanation of the effect suggests that a low level of cross-race contact leads to an impoverished ability to differentiate, using the appropriate cues, the faces of other-race individuals (e.g. Meissner & Brigham, 2001; Sporer, 2001). Given this theoretical position, at least one key issue might be the psychological construal of other-race faces, which might be affected by the amount of contact the witness has had with other-race faces. An expert, of course, is not able to test the eyewitness in an actual case to know the level of experience the witness has had with other-race faces. Still further, there is no research to tell us what degree of experience a witness should have with other-race faces before accuracy is expected to improve, nor do we have any empirical basis for recommending how “other-race face experience” ought to be reliably measured in actual witnesses.

A final point we wish to touch on is that in testifying in the courtroom, there is the possibility that an expert will confound constructs with levels of constructs. As an example, assume that an expert testifies that eyewitnesses have more difficulty remembering stressful than nonstressful events. More aptly put, research finds that memory (as studied in a particular context, measured in a particular way, and with a particular type of participant) performance (as measured with a particular type of memory test) decreases given a certain level (and type) of exposure to stress compared to some certain baseline level of nonstress. Exposure to stress per se does not change memory, but rather some degree (or quality) of stress changes behavior which in turn may affect memory. For instance, we would not automatically conclude that a student will perform well on an exam because he has studied, or that bombarding a plutonium atom will cause a nuclear explosion. In all cases, we have to take into account the dosage level of the precipitating factor and the other variables that may have been present at the time of witnessing in order to postdict whether the factor may have negatively impacted memory.

Background conditions may affect causal relationships

Meta-analyses of lab studies (e.g., Bradfield-Douglas & Steblay, 2006; Steblay, 1992) are often done with an eye toward making a particular research area more amenable to courtroom presentation. A typical conclusion reached from this work is that a variable is an important one and reliable enough for generalization to the legal system if its effect is statistically significant in a majority of the studies that has investigated the phenomenon. To illustrate, Bradfield-Douglas and Steblay (2006), performed a meta-analysis of the post-identification feedback effect, and concluded based on their findings: “For experts who testify in court, this meta-analysis will facilitate admittance of testimony on this topic... The consistency in outcomes demonstrated in this review lends credence to the argument that post-identification feedback effects should be

‘generally accepted’” (pp. 865-866). The range of conditions in which an effect has been investigated, however, may have a direct relationship to the range of population characteristics to which generalization is possible (Meehl, 1989). The closer a researcher’s enterprise is to a domain of application, the more important it is to raise questions regarding external validity. This point is made well by Mook (1983): “Of course there are also those cases in which one does want to predict real-life behavior directly from research findings. Survey research, and most experiments in applied settings such as the factory or classroom, have that end in view. Predicting real-life behavior is a perfectly legitimate and honorable way to use research. When we engage in it, we do confront the problem of E[xternal] V[alidity]...” (p. 386).

One might contend that there is sufficient variability across studies in the operationalization of key variables and in the background conditions employed (i.e., the variables that are held constant across the experimental and control conditions, such as duration of exposure to the culprit, the lineup employed, the to-be-remembered event), and as such, we have explored a sufficient range of conditions for many variables (e.g., weapon exposure, cross-race witnesses, stress) to be able to broadly generalize the cause and effect relationships observed across laboratory studies to specific cases in the legal system. In this section, we compare the similarity of laboratory and field eyewitness contexts, and the results suggest that we have charted only a part of the distance.

Table 1 presents the background conditions of papers summarized by well-known meta-analyses (Meissner & Brigham, 2001; Steblay, 1992; Steblay et al., 2001). Of particular interest here are the background conditions employed by these studies that affect memory strength, namely, the length of the critical event exposure duration and the length of the retention interval between the to-be-remembered event and the identification task. If the meta-analysis study

authors did not provide summary information about memory strength, we obtained the information from the original published reports that were part of the meta-analysis (i.e., “grey literature”, or unpublished work, was not included our analysis). In Table 2 we present the average duration of exposure and retention interval found in our analysis of 721 cases (rape, robbery, and assault) involving lineup identifications that were randomly selected from files that were submitted by the San Diego Police Department to the San Diego District Attorney’s Office for review. As a comparison of Tables 1 and 2 indicates, the background conditions in the laboratory studies differ from the conditions experienced by archival eyewitnesses making identifications. The duration of exposure to the culprit and the identification retention interval typically employed by the laboratory studies are short relative to the archives.

Our analysis of duration of exposure to the culprit was the result of estimating the duration on the basis of the information reported in the case file (including witness statements in the majority of the cases, police incident reports, etc.). Therefore, because the validity of our estimates can be questioned, we also analyzed a field study (Woolnough & MacLeod, 2001) that examined the accuracy of testimony given by witnesses in real world cases by comparing the contents of the witness reports to closed circuit television video recordings of the witnessed crime. The average duration of exposure to the culprit (based on our analysis of the raw data presented in the report) was 1 minute and 27 seconds (median: 1 minute and 20 seconds; range: 20 seconds to 3 minutes and 45 seconds), which again is relatively long compared to the laboratory studies.

Yet another crucial difference between the laboratory research and the field is that eyewitnesses in the field may be questioned on multiple occasions, whereas laboratory witnesses are typically questioned once (Ebbesen & Rienick, 1998). Accuracy is expected to be lower for eyewitnesses who render testimony after a long compared to a short delay (e.g. Shapiro &

Penrod, 1986). Real world eyewitnesses, however, experience retention intervals differently than eyewitnesses in the lab if they provide testimony at several points in time; usually they are questioned immediately after the event, and then yet again at other points in time (Civilini & Flowe, 2008). Ebbesen and Rienick (1998) demonstrated, by comparing memory loss functions within and between subjects, that forgetting did not occur if participants were questioned repeatedly within the retention interval. This work illustrates that it is crucial to time experimental procedures in a manner that is comparable with how police question and test eyewitnesses. Experts testifying in the courtroom assume that the effects obtained in the laboratory generalize across memory strength conditions. It is highly plausible, however, that many of the effects that experts testify about are moderated by memory strength.

Even if we did have comparable levels of memory strength in the lab and in the real world, we still might wonder whether the real world and laboratory contexts are comparable in other respects. We located 290 published experiments on lineup/showup identification (conducted from 1975-2006) from the PsychInfo database. Of the studies, 44% presented the target in a video, 17% live in a lab, 13% live in an auditorium or a classroom, 11% in a photograph, 8% live in a natural setting, and 7% in a slide sequence. The target was portrayed under conditions that were not criminal in 41% of the studies (e.g., laboratory participants memorized a photograph, or field participants were asked to identify a customer or a researcher with whom they had previously interacted), and the target committed theft in 32%, robbery in 20%, other types of crimes in 4%, assault in 1%, and committed rape in less than 1% of the studies. Additionally, the majority of studies involved college students (68% of the studies recruited exclusively college students, 23% were from other adult populations, and 9% were children). These results indicate that the eyewitness memory studies most commonly study (based on the modes presented above)

noncriminal activities presented in videotapes to college students. The research context employed may of course vary across the domain of investigation (e.g., lineup research, misinformation effect research, stress research). However, one might still wonder whether differences between the contexts in which the research is carried and the contexts in which eye witnessing occurs are important considerations for determining whether an experimental effect generalizes.

One response to our contention that laboratory and real world contexts may not be comparable is the argument that memory for real world crime events is bound to be worse compared to memory for laboratory events. This point is often raised in arguing that the laboratory studies are generalizable because they represent the best case scenario: “The results showed that participants who watched the videos reported more details and with higher accuracy than those who saw the live events, suggesting that laboratory experiments may actually overestimate memory performance” (Loftus, 2003, p. 868). Though this could be true, we simply do not know. How background factors (such as memory strength, or eyewitness vs. student witness/lab participant motivation and intelligence) might interact with primary eyewitness variables (such as weapon focus, race of the eyewitness and perpetrator, or stress) is an empirical issue that necessitates further investigation. Additionally, the results of field studies raise the possibility that variables (such as memory strength) within the eyewitness context might moderate the effects of primary eyewitness variables. For instance, field studies examining the weapon focus effect did not find significant differences in suspect identification rates depending on whether a weapon was present (Behrman & Davey, 2001; Valentine, Pickering, & Darling, 2003; Wright & McDaid, 1996), or depending on whether the witness was subjected to violence (Wright & McDaid, 1996). Additionally, field studies have found mixed results regarding the

cross-race effect (Behrman & Davey, 2001; Valentine, Pickering, & Darling, 2003). As another example, Yuille and Cutshall (1986) found that stress and accuracy were not associated in their case study of statements made by actual eyewitnesses to a shootout. They further found that witnesses who reported higher levels of stress were closer in proximity to the perpetrator than witnesses reporting lower levels of stress. In other words, in actual cases, stress and vantage point of the witness are likely confounded. Even though archival studies are correlational and cause and effect relationships cannot be deduced from the results, they draw our attention to the complexity of circumstances that exist outside of the laboratory that may serve to moderate the main effects of variables on accuracy that are found in the laboratory.

The incremental validity of using a given factor to detect errors in actual cases depends on more than the effect size of a given variable

A key issue that is addressed in other domains of applied psychology (e.g., organizational psychology, clinical psychology, health psychology, personality psychology) is *incremental validity*, which refers to the extent to which a particular selection procedure improves on decision making in the context of application. Incremental validity is important because the validity of a selection test can vary depending upon the context of application (see Hunsley & Meyer, 2003 for further commentary on issues concerning incremental validity in applied psychology). Selection tests can be costly, both financially and in human terms; hence, it is important to assess whether the goal behind using a particular selection test is achieved in a particular context.

In testifying in the courtroom, we are assuming that expert testimony on the issue of eyewitness memory, which is selection procedure for identifying cases that are possibly problematic, has incremental validity over not having the testimony. Traditional safe guards,

such as police and prosecutor case screening procedures and cautionary instructions to the jury regarding the factors that can affect the reliability of eyewitness testimony (see Devenport, Kimbrough, & Cutler, Chapter 3, this volume, for a review), have been argued to be ineffective in preventing erroneous convictions based on erroneous eyewitness testimony. As such, expert testimony about risk factors that decrease the accuracy of eyewitness testimony may be advisable.

The size of the effect, or the validity coefficient for the construct about which the expert is testifying (r) (see Pezdek, Chapter 2, this volume), is but one factor to consider in determining whether expert testimony is warranted. Meta-analytic results indicate that the effect size estimate for weapon focus in lineups ($r = .09$) (Stebly, 1992), for instance, is smaller than the effect size estimate for biased instructions in target absent lineups ($r = .38$) (Cutler & Penrod, 1995). We could conclude based on this comparison that the biased instruction effect is relatively more robust than the weapon focus effect, and therefore, is a factor about which we should testify more often. However, there are other factors that have to be taken into account as well in determining what is gained by providing the expert testimony to the courts. Incremental validity in applied psychology is examined by taking into account the effect size of the predictor, the base rate of the outcome, the selection ratio, or the number of cases that will be examined in the applied context, and the definition of the construct in question. In the eyewitness memory domain, the incremental validity of utilizing testimony in a given jurisdiction would be assessed by determining:

- 1) the effect size of the predictor
- 2) the base rate of innocent versus guilty defendants
- 3) the selection ratio, or the proportion of cases in which the expert testifies; and

4) the applied definition of the construct (i.e., the type of scenarios the expert qualifies as meeting the definition of a construct, such as what types of scenarios would qualify as involving weapon focus or mug shot induced bias)

To illustrate, consider a scenario in which the effect size for a risk factor is $r = .40$ and the selection ratio is 20%. We can examine the *true positive rate*, or the proportion of cases in which the expert would testify when the defendant is in fact innocent. We can also examine in relation to the base rate of innocence the *false positive rate*, or the proportion of cases in which the expert would testify when the defendant is in fact guilty. Holding the effect size and selection ratio constant, we can compute the true positive rate and the false positive rate as a function of different base rates of innocence.¹ The results of the analysis are shown in Figure 1. As can be seen, if 15% of defendants appearing in the courtroom are in fact innocent, then the true positive rate will be 3%, and the false positive rate will be 17%. In other words, in a population of 100 cases in which 15 defendants are actually innocent, the expert would provide testimony in 3 cases. For the remaining 85 defendants who are actually guilty, the expert would testify in 17 cases. Under the conditions described, the true positive rate will equal or exceed the false positive rate only when the base rate of innocence is equal to or greater than 50%. This analysis makes clear that there is the risk of false positives when expert testimony is utilized.

Though not central to the discussion, the false negative and true negative rates refer to occasions when the expert does not testify. The *true negative rate* (the proportion of cases in which the expert does not testify and the defendant is guilty) and the *false negative rate* (the proportion of cases in which the expert does not testify and defendant is innocent) and are also provided in Figure 1. As can be seen, if the base rate of innocence is 15% and the selection ratio is 20%, out of 100 cases there will be 68 in which the defendant is guilty (true positive rate), and

12 in which the defendant is innocent (false negative rate). One might think that the expert should testify more often to decrease the false negative rate. Doing so, however, would simultaneously increase the false positive rate.

Lastly, the type of scenarios that the expert qualifies as weapon focus is also important. The size of the effect can be expected to vary depending on the standards employed. For instance, what type of instrument (e.g., gun, knife, screwdriver) qualifies as a weapon (see Pickel, 1998), and for how long should the witness be exposed to the perpetrator in order for the weapon focus effect to hold (see Steblay, 1992 for meta-analysis of moderators of weapon focus effect in mostly target present lineups)?

As a means for further discussing some of the difficulties an expert might encounter in mapping laboratory variables onto real world scenarios, consider the case of nineteen year old Adam Noriega, who was accused of shooting a rival gang member to death (People v. Noriega, 2003). Eyewitness identifications of Noriega served as the primary evidence and were made by his companion that was present during the shooting and by D.S., a fourteen year old witness who saw the shooting take place from across the street. The expert in the case presented a written motion to the court regarding the possible negative effects of “photobiasing” on eyewitness identification accuracy. The eyewitness expert explained to the court that “photobiasing” is a priming effect whereby a witness identifies a suspect from the lineup because he or she had been previously exposed to the suspect’s photograph. D.S. had been presented with 30 photographs of local area gang members, one of which was a photograph of Noriega taken at a younger age, and she did not identify anyone. Six weeks later she positively identified a recent photograph of Noriega in a lineup test. The expert was ultimately not allowed to testify because the photograph of Noriega taken at a younger age bore little resemblance to him at the time of the shooting, an

observation to which the defense conceded. The exclusion of the expert testimony was later raised on appeal, however, after Noriega was found guilty of premeditated murder. For our present purposes, this case raises interesting issues concerning how to translate effects found in laboratory studies to actual cases. The term given in the research literature to the biasing effect of previously having been exposed to mug shots on lineup identification accuracy is called *mug shot induced bias* (see Kassin et al., 2001 survey). Some of the questions that arise in mapping mug shot bias research onto the Noriega case include: If a witness does not identify a suspect from a photo that does not look like him, and later the witness identifies the suspect from a photo that does look like him, is the positive identification an example of mug shot induced bias? Does mug shot induced bias arise in the absence of a positive identification from the mug shots? (Research by Dysart and colleagues (2001), which may not have been available at the time of Noriega's trial, suggests that it does not.) What is the relationship between the age of a suspect in a photograph and the suspect's current age with respect to eyewitness identification accuracy, a topic about which there is no research?

The foregoing example illustrates that there is considerable room for judgment on the part of the expert in deciding whether a particular effect found in a laboratory study applies to a real world scenario. Expert judgment in this translation has a direct effect on the incremental validity of expert testimony. It is unknown how often the issues we raise present difficulty for experts. To be sure, we do not know how often some of the translational difficulties found in the *People v. Noriega* (2003) case arise in the cases that experts retain, as systematic research has not been done on the issue. If these difficulties commonly occur, then it would be interesting to find out how experts go about solving them. There may be considerable agreement across experts that certain variables have reliable effects in the laboratory and that the effects are reliable enough on

which to testify (Kassin et al, 2001). But whether there is agreement across experts in the application of the definitions of these variables in real world cases is an empirical issue that remains to be addressed.

Eyewitness selection biases in the field may limit the comparability of real life and laboratory eyewitnesses

Some who advocate applying laboratory research in the courtroom might argue that laboratory eyewitnesses have better memories, better visual acuity, and are more alert than are other adult populations (e.g., Wells, 2004). A conclusion that could be reached based on this supposition is that accuracy rates are bound to be higher for laboratory compared to real world eyewitnesses (e.g., Loftus, 2003).

Though factors such as memory ability surely affect memory accuracy, we simply do not know whether the eyewitnesses selected by the police to testify differ in memory ability and motivation compared to laboratory eyewitnesses. We know based on our analysis of police cases forwarded to the Prosecutor's Office in San Diego that on average only about 50% of the witnesses in a case are given a lineup test. In contrast, laboratory studies administer a lineup test to every participant, without regard to motivation, vantage point, or memory strength of the participant. If witnesses in real world cases are selected on the basis of these factors—and these factors moderate cause and effect relationships demonstrated in the lab, then generalizations from the laboratory to the courtroom are questionable. Moreover, witnesses may self-select into identification groups, and decline to take an identification test if their memory is weak: “A true weapon-focus effect could be obscured if witnesses to crimes involving weapons believe that their memory is weak and are therefore less inclined to attend lineups. The result could be a reduction in the number of weapon-focus-impaired witnesses presented with lineups and thus a

reduced number of cases of weapon focus” (Wells, Memon, & Penrod, 2006, p. 53). Still further, a weapon focus effect may not be obtained in field studies because witnesses exposed to weapons agree (or are asked by the police) to take a lineup test only if they attended to the perpetrator’s appearance, whereas witnesses who did not attend to the perpetrator’s appearance, decline (or are not asked to attempt an identification). In short, if such selection biases operate in the legal system, research participants may be unrepresentative of the characteristics of witnesses in actual criminal cases.

Experts Should Not Generalize From DNA Exoneration Cases

Eyewitness identification evidence plays a crucial role in apprehending the guilty. The following example attests to the power and accuracy of human memory:

On October 2nd, 1975, three eyewitnesses positively identified Theodore Robert Bundy from a seven-person lineup. One of the eyewitnesses was Carol DaRonch, who indicated that Bundy had kidnapped her, threatened her with a handgun, and attempted to smash her head in with a crowbar before she managed to fight him off and escape his clutches. Although Bundy repeatedly professed his innocence, the police launched a full-blown investigation after Bundy was positively identified.

Of course, this example does not prove the aforementioned premise. This anecdote says nothing at all regarding the rate at which eyewitnesses correctly identify guilty suspects, nor does it tell us why the witnesses in the case might have been correct. Still further, this example draws our attention away from the fact that witnesses can also fail to identify a guilty perpetrator, who in turn kills other people, or that witnesses can mistakenly identify an innocent person, who in turn languishes in prison while the guilty perpetrator is free to kill.

Despite the inherent limitations of anecdotal evidence, researchers in the eyewitness memory domain often introduce their work by referring to actual cases of mistaken identification. In fact, so far this year (2007), nearly a third of the published research papers on eyewitness identification utilize anecdotes of mistaken identification cases to introduce their topic of inquiry

(Busey & Loftus, 2007; Haw, Dickinson, & Meissner, 2007; Keast, Brewer, & Wells, 2007; Krug, 2007; Lindsay, 2007; MacLin, & Phelan, 2007; Neuschatz, Lawson, Fairless, Powers, Neuschatz, & Goodsell, 2007; Remijn, & Crombag, 2007; Wells & Hasel, 2007). Additionally, DNA exoneration cases in relation to the issue of mistaken eyewitness identification are also cited in cases in which eyewitness memory experts have been involved (e.g., *People v. Adams*, 2008; *People v. Copeland*, 2007; *United States v. Burton*, 1998).

One problem with drawing conclusions from a specific case regarding the factors that may increase the likelihood of mistaken identification is that the characteristics of the case may be unrepresentative of false identifications in general. For instance, the fact that the majority of DNA exoneration cases involve sexual assault does not mean that eyewitness memories are particularly weak in sexual assault cases; rather, sexual assault cases have DNA evidence available for analysis, whereas other cases do not, and therefore, potential eyewitness errors may be identified in sexual assault cases but not in other cases like robbery (Wells, Memon, & Penrod, 2006). Nevertheless, despite the fact that DNA exoneration cases are not randomly sampled, and hence unrepresentative of false identifications in general, experts providing information to the court may draw a parallel between eyewitness factors present in the DNA exoneration cases and similar factors that were present in the case at hand. The intention in so doing is to argue that eyewitnesses who are exposed to factors like those found in DNA exoneration cases tend to be inaccurate. However, in order to determine whether a given factor is influential, the sample should also include cases in which eyewitnesses made correct identifications (see Finklea and Ebbesen, 2007, for such an analysis).

The DNA exoneration cases appear to have had a profound influence on the admission of expert testimony on eyewitness identification. Because the DNA exoneration cases seem to often

arise in court discussions (People v. Dubose, 2005; People v. Herrera, 2006; People v. Lawrence, 2007; People v. Lewis, 2008; People v. Romero, 2007; People v. Williams, 2006; People v. Woolcock, 2005), we feel that it is worth going into a bit more detail about the extent to which these cases are informative with regard to using them in other case to weigh the potential for eyewitness inaccuracy. Predicting the contexts in which mistaken identifications are likely to occur requires analyzing a broad spectrum of cases, not just the cases in which eyewitnesses have made "known" errors. Observing a common feature among exoneration cases does not necessarily mean that the feature caused the mistaken identifications to occur. The observed feature might be common even among cases in which the eyewitness identifications were accurate. Additionally, the frequency of the causal feature in question, such as cross-race eyewitnesses, may fluctuate across the different levels of the criminal justice system. To illustrate, Gross et al. (2005) argued that cross-race identifications played a major role in the DNA exoneration cases. This proposition is difficult to evaluate, however, in the absence of knowledge regarding the distribution of cross-race cases across the various levels of the criminal justice system (e.g., arrest, prosecution, convictions levels). Gross et al. (2005) found that in 69 DNA exoneration cases in which black males had been convicted of raping a stranger, just about 50% involved white victims. They reasoned that the number of white victims among the black exoneration cases should have instead been on the order of 5-6%, because victim surveys indicate that most perpetrators are white men (90%), that most rapes are committed within race (88%), and that the rate of being a survivor of rape is equal for black and white women. Thus, since systematic research finds that people are prone to making cross-race identification errors, they concluded that the overrepresentation of black defendants among the exoneration rape cases suggests that mistaken identifications were caused by the cross-race effect.

One might conclude on the basis of Gross et al.'s (2005) analysis that the DNA exoneration cases demonstrate that cross race identifications are more likely to result in eyewitness error. However, there is another alternative explanation for their findings which derives from an analysis of case flow in the criminal justice system. The base rate of interracial rape varies across the different levels of the criminal justice system. The base rate of black on white victimizations is 27% among stranger cases reported to the police (U.S. Department of Justice, 1992-2004). Gross et al. (2005) derived their estimate of interracial rape based on the characteristics of National Crime Victimization Survey respondents who reported having experienced rape, and these respondents may not be representative of the characteristics of the victims/defendants in rape cases reported to the police, or in rape cases that are prosecuted, or in rape cases in which there is a conviction. The vast majority of women indicating on victim surveys that they were raped do not report the rape to law enforcement, especially rape perpetrated by acquaintances. Additionally, in stranger rapes, both black and white victims are more likely to report rape to the police when the assailant is black. This contention is supported by the fact that although victim surveys indicate that 90% of the assailants are white men, the percentage of black (35%) and white (37%) men in prison for rape is approximately equal (U.S. Department of Justice, 2005). Still further, even though rape is more likely to occur within race, black males might more likely to be convicted for raping white compared to black women. Research with mock jurors has found that the probability of a guilty verdict is reduced in cases that involve black women as complainants (Klein & Creech, 1982). Therefore, the percentage of stranger rape convictions involving black defendants and white victims most certainly exceeds 27% (i.e., the rate of reporting black on white stranger rape to the authorities). As a result, when black men are convicted of raping strangers, there will be a large number of cases involving white women.

Thus, Gross et al.'s (2005) contention that the disproportionate number of black on white rape exoneration cases is highly suggestive that the cross race identifications is led to the wrongful convictions can be alternatively explained by differential rape reporting and prosecution rates stemming from the race of the victim and defendant. Indeed, in the random sample of rape cases that were referred by the police for prosecution in San Diego, across rape survivors who were previously unacquainted with the perpetrator, 47% were cross-race.

By the foregoing analysis, we are not arguing that the cross-race effect plays no role in wrongful convictions. Rather, our point is that the expected frequency of the feature in question might not be equal across the different levels of the criminal justice system due to case selection effects. Therefore, we have to be cautious in drawing conclusions if we find that certain factors are common in exoneration cases. The common factors may not be necessarily causal with regard to mistaken identification.

Moving Forward: Summary and Conclusions

In this chapter, we have argued that generalizing from the eyewitness identification literature to specific cases in the legal system by simply matching variables across laboratory and real world contexts may be invalid. Variables in the laboratory are inextricably tied to the totality of the specific procedures that were used to instantiate the variable. If the procedures that were used to instantiate the variable in the laboratory differ from the operations that give rise to the variable in the context of the legal system, then the nomological network (Cronbach & Meehl, 1955) in the lab and in the real world may very well be different. In other words, if a variable has a causal effect on behavior in the laboratory, a similar-looking construct that presents in a real world case may not have the same effect on behavior in the real world as it does in the laboratory. We also reasoned in this chapter that the unqualified application of main effects research to a particular

case in the legal system may too be invalid. The dimensions on which laboratory and real world contexts differ is important because of the fact that there may be factors present in real world cases, such as memory strength, that serve to moderate the negative impact of a variable that adversely affects memory in the laboratory. We also argued that courtroom testimony that includes discussion of DNA exoneration cases is invalid because we cannot infer anything whatsoever from DNA exoneration cases alone regarding the factors that cause mistaken identifications.

We believe that the best application of our science to the legal system is field tests that enable us to compare procedures for collecting and preserving eyewitness identification evidence. Our science does not permit us to make predictions regarding whether a given factor has affected a particular eyewitness' testimony. We can, however, determine how a particular procedure, implemented in a particular manner, in a particular context influences suspect identification rates in the long run. We cannot expect that courtroom testimony in the long run will prevent more errors than not, due the difficulties inherent in having an expert map laboratory factors onto a given case (see Dawes, 1994 for a discussion of clinical versus actuarial prediction). Still further, as illustrated in this chapter, the incremental validity of expert testimony will vary depending on the base rate of defendant innocence. We run the risk of presenting to fact finders information regarding the factors that negatively impact eyewitnesses in cases in which the defendant is actually guilty. Moreover, by analyzing the natural context in which eyewitnesses render testimony, we can improve the representativeness (see Hammond & Stewart, 2001, for an in depth treatment of Egon Brunswik's conception of representative design) of our experimental designs and thereby increase their applicability.

The external validity of our research is an important empirical issue, much too important to leave up to chance; it is not enough to assume that after 20 years of conducting research in a particular field that the background conditions will vary widely enough across studies to permit the generality of the findings, or to assume, without empirical evidence from the field, that real world conditions are more likely to produce inaccurate identifications than laboratory studies (e.g., Loftus, 2003). We hope that the points we have raised in this chapter will invigorate interest in describing the natural context in which eyewitness identifications are carried out, and stimulate researchers to empirically delineate the real world contexts to which a theory applies. Finally, prospective field studies are needed in order to test within context procedural recommendations that have been proposed for increasing the reliability of eyewitness testimony.

References

- Agresti, A. (2002). *Categorical data analysis*. New York: Wiley.
- Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, 5(1), 1-21.
- Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and human behavior*, 25(5), 475-491.
- Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, 37(3), 245-257.
- Bradfield-Douglass, A. B., & Steblay, N. (2006). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology*, 20(7), 859-869.
- Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology*, 87, 112-120.
- Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own- and other-race persons. *Journal of Social Issues*, 41(3), 139-155.
- Busey, T. A., & Loftus, G. R. (2007). Cognitive science and the law. *Trends in cognitive sciences*, 11(3), 111-117.
- Civilini M. C., & Flowe, H. D. (2008). *An archival analysis of the factors that affect the completeness of eyewitness statements*. Poster presentation at the Western Psychological Association meeting, Irvine, CA.
- Clifford, B.R., & Lloyd-Bostock, S.M.A. (1983). Witness evidence: Conclusions and prospect.

- In S. M. A. Lloyd-Bostock, & B. R. Clifford (Eds.), *Evaluating witness evidence: Recent psychological research and new perspectives*. New York: Wiley.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
- Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology and law*. New York: Cambridge University Press.
- Danziger, K., & Dzinas, K. (1997). How psychology got its variables, *Canadian Psychology*, 38, 43-48.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. The Free Press, New York, NY.
- Deffenbacher, K. A. (1984). Experimental psychology actually can assist triers of fact. *American Psychologist*. 39(9), 1066-1068.
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, 28, 687-706.
- Ebbesen, E. B., & Konečni, V. J. (1997). Eyewitness memory research: Probative v. prejudicial value. *Expert Evidence: The International Digest of Human Behaviour, Science, and the Law*, 5, 2-28.
- Ebbesen, E. B., & Rienick, C. B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology*, 83(5), 745-762.
- Elliott, R. (1993). Expert testimony about eyewitness identification: A critique. *Law and Human Behavior*, 17(4), 423-437.
- Fahsing, I.A., Ask, K. & Granhag, P.A. (2004). The man behind the mask: Accuracy and

- predictors of eyewitness offender descriptions. *Journal of Applied Psychology*, 89, 722-729.
- Feingold, G. A. (1914). The influence of environment on identification of persons and things. *Journal of Criminal Law & Criminology*, 5, 39-51.
- Finklea, K. M., & Ebbesen, E. B. (2007). Eyewitness accuracy in the real world: DNA evidence as “ground truth” for eyewitness accuracy rates. *Unpublished manuscript*.
- Gross, S. R., Jacoby, K., Matheson, D. J., Montgomery, N., & Patel, S. (2005). Exonerations in the United States, 1989-2004. *The Journal of Criminal Law & Criminology*, 95(2), 523-560.
- Haber, R. N., & Haber, L. (2000). Experiencing, remembering, and reporting events. *Psychology, Public Policy, and Law*, 6(4), 1057-1097
- Hammond, K. R., & Stewart, T. R. (Eds.). (2001). *The essential Brunswik: Beginnings, explications, applications*. New York, NY, US: Oxford University Press.
- Haw, R. M., Dickinson, J. J., & Meissner, C. A. (2007). The phenomenology of carryover effects between show-up and line-up identification. *Memory*, 15(1), 117-127.
- Hope, L., Ost, J., Gabbert, F., Healey, S., & Lenton, E. (2008). “With a little help from my friends...”: The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica* 127(2): 476-484.
- Hunsley, J., & Meyer, G.J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, 15, 446-455.
- Keast, A., Brewer, N., & Wells, G. L. (2007). Children's metacognitive judgments in an eyewitness identification task. *Journal of Experimental Child Psychology*, 97(4), 286-314.
- Klein, K., & Creech, B. (1982). Race, rape, and bias: Distortion of prior odds and meaning changes. *Basic and Applied Social Psychology*, 3(1), 21-33.

- Konečni, V. J., & Ebbesen, E. B. (1986). Courtroom testimony by psychologists on eyewitness identification issues: Critical notes and reflections. *Law and Human Behavior*. Special Issue: The ethics of expert testimony, 10(1-2), 117-126.
- Krug, K. (2007). The relationship between confidence and accuracy: Current thoughts of the literature and a new area of research. *Applied Psychology in Criminal Justice*, 3(1), 7-41.
- Lindsay, D. S. (2007). Autobiographical memory, eyewitness reports, and public policy. *Psychologie Canadienne*, 48(2), 57-66.
- Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior*, 18, 527-541.
- Loftus, E.F. (1983). Silence is not golden. *American Psychologist*, 38, 564-572.
- Loftus, E. F. (2003). Make-believe memories. *American Psychologist*, 58(11), 867-873.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning & Memory*, 4(1), 19-31.
- MacLin, O. H., & Phelan, C. M. (2007). PC_Eyewitness: Evaluating the New Jersey method. *Behavior Research Methods*, 39(2), 242-247.
- McCloskey, M., & Egeth, H. E. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist*, 38(5), 550-563.
- McCloskey, M., Egeth, H., & McKenna, J. (1986). The experimental psychologist in court: The ethics of expert testimony. *Law and Human Behavior*. Special Issue: The ethics of expert testimony, 10(1-2), 1-13.
- McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for

- events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, 114(1), 1-16.
- Meehl, P. E. (1989). Law and the fireside inductions (with Postscript): Some reflections of a clinical psychologist. *Behavioral Sciences and the Law*, 7, 521-550.
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3-35.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38(4), 379-387.
- Neuschatz, J. S., Lawson, D. S., Fairless, A. H., Powers, R. A., Neuschatz, J. S., & Goodsell, C. et al. (2007). The mitigating effects of suspicion on post-identification feedback and on retrospective eyewitness memory. *Law and Human Behavior*, 31(3), 231-247.
- People v. Adams, No. A-103, 2008, N.J. LEXIS 144.
- People v. Copeland, No. E2002-01123-SC-DDT-DD, 2007, Tenn. LEXIS 502.
- People v. Davis, No. D042297, 2004, C.A. LEXIS 4369.
- People v. Dubose, No. 2003AP1690-CR, 2005, W.I. LEXIS 400.
- People v. Herrera, 187 N.J. 493; 902 A.2d 177; 2006, N.J. LEXIS 1047.
- People v. Ellison, No. D048918, 2008, C.A. LEXIS 1292.
- People v. Garcia, B195201, 2008, C.A., LEXIS 6665.
- People v. Lawrence, SC 17452, Conn., 2007, LEXIS 172.
- People v. Lewis, NY Slip Op 51747U; 20 Misc. 3d 1136A; 2008 N.Y. Misc. LEXIS 4873.
- People v. Noriega, B152427; 2003, C.A. LEXIS 9087.
- People v. Romero, 191 N.J. 59; 922 A.2d 693; 2007 N.J. LEXIS 585.
- People v. Williams, Slip Op 26469; 14 Misc. 3d 571; 830 N.Y.S.2d 452; 2006 N.Y. Misc. LEXIS 3429.

People v. Woolcock, NY Slip Op 25045; 7 Misc. 3d 203; 792 N.Y.S.2d 804; 2005 N.Y. Misc.

LEXIS 157.

Pickel, K.L. (1998). Unusualness and threat as possible causes of "weapon focus." *Memory*, 6, 277-295.

Remijn, C. A. C., & Crombag, H. F. M. (2007). Heuristics in causal reasoning and their influence on eyewitness testimony. *Psychology, Crime & Law*, 13(2), 201-211.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139-156.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, 7(1), 36-97.

Sporer, S. L. (1996). Describing others: Psychological issues. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 53-86). Hillsdale, NJ: Lawrence Erlbaum.

Stebly, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, 16(4), 413-424.

Stebly, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, 25(5), 459-473.

United States v. Burton, No. 3:97-cr-154, 1998, U.S. Dist. LEXIS 18730.

U. S. Department of Justice (1992-2004). *National Crime Victimization Survey*. Washington, D.C.: U.S. Department of Justice.

U. S. Department of Justice (2005). *Uniform Crime Reports*. Washington, D.C.: U.S. Department of Justice.

- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology, 17*(8), 969-993.
- Windschitl, P. D. (1996). Memory for faces: Evidence of retrieval-based impairment. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*(5), 1101-1122.
- Wells, G. L. (2004). *Primer on eyewitness ID research*. Retrieved April 14, 2008 from http://www.nlada.org/Defender/forensics/for_lib/Index/Eyewitness%20ID/Eyewitness%20ID%20Defense%20Expert%20Materials/Prof.%20Gary%20Wells/index_html#Prof.%20Gary%20Wells
- Wells, G. L., & Hasel, L. E. (2007). Facial composite production by eyewitnesses. *Current Directions in Psychological Science, 16*(1), 6-10.
- Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7*(2), 45-75.
- Woolnough, P. S., & MacLeod, M. D. (2001). Watching the birdie watching you: Eyewitness memory for actions using CCTV recording of actual crimes. *Applied Cognitive Psychology, 15*(4), 395-411.
- Wright, D. B., & McDaid, A. T. (1996). Comparing system and estimator variables using data from real line-ups. *Applied Cognitive Psychology, 10*(1), 75-84.
- Wright, D. B., & Stroud, J. N. (1998). Memory quality and misinformation for peripheral and central objects. *Legal and Criminological Psychology, 3*, 273-286.
- Yarmey, A. D. (1997). Probative v. prejudicial value of eyewitness memory research. *Expert Evidence, 5*(3), 89-97.
- Yarmey, A. D. (2001). Expert testimony: Does eyewitness memory research have probative value for the courts? *Canadian Psychology/Psychologie Canadienne, 42*(2), 92-100.

Yuille, J. C. (1993). We must study forensic eyewitnesses to know about them. *American Psychologist*, 48(5), 572-573.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71(2), 291-301.

Footnote

¹ The values in Figure 1 were calculated by assuming a 2 X 2 contingency table, where a = true positives, b = false positives, c = false negatives, and d = true negatives. The following formulae were used to solve for these values (following Agresti, 2002, and advice obtained in personal communication with James H. Derzon, from The Centers for Public Health Research and Evaluation, Battelle):

$$a = N(p_{r1}P_{c1} + r\sqrt{p_{r1}r_{ci}(1-p_{r1})(1-p_{c1})})$$

$$b = Np_{r1} - a$$

$$c = Np_{c1} - a$$

$$d = N - (a + b + c)$$

where r = the correlation coefficient for a 2X2 contingency table, N = Total sample size, p_{r1} = the selection rate for the predictor, and p_{c1} is the base rate for the outcome.

Table 1.

Critical event exposure duration and identification retention interval by research literature.

	Critical Event Exposure Duration		Retention Interval	
	<i>Median</i>	<i>Range</i>	<i>Median</i>	<i>Range</i>
Weapon Focus Effect	45 s	18 s - 240 s	20 min	immediate - 36 days
Cross-Race Effect	3 s	.12 s - 4 min	2 min	immediate - 21 days
Simultaneous and Sequential	1.25 min	4.5 s- 6 min	15 min	immediate - 7 days

Table 2.

Archival results for critical event exposure duration and identification retention interval by case type.

	Critical Event Exposure Duration		Retention Interval	
	<i>Median</i>	<i>Range</i>	<i>Median</i>	<i>Range</i>
Assault	10 min	4.8 sec to 9 hours	18 days	1 day - 90 days
Rape	10 min	30 sec to 120 hours	2.97 days	1 day - 72 days
Robbery	5 min	15 sec to 7.5 hours	5.31 days	1 day - 44 days

Figure 1.

Hypothetical example illustrating the effects of base rates of defendant innocence and expert testimony on case outcomes. The example assumes that the effect size for the risk factor about which the expert testifies is $r = .40$, and the rate at which the expert testifies in cases (i.e., the selection ratio) is $.20$. The *true positive rate* reflects the proportion of cases in which the expert testifies when the defendant is innocent. The *false positive rate* is the proportion of cases in which the expert testifies and the defendant is guilty. The *false negative rate* indicates the proportion of cases in which the expert does not testify and the defendant is innocent, and the *true negative rate* indicates the proportion of cases in which the expert does not testify and the defendant is guilty.

Figure 1

