ANALYSIS OF SCIENTIFIC RESEARCH ON EYEWITNESS IDENTIFICATION

January 2018

Patricia A. Riley
U.S. Attorney's Office, DC

Table of Contents

SUMMARY

RESEARCH DOES NOT PROVIDE A FIRM FOUNDATION FOR JURY INSTRUCTIONS OF THE TYPE ADOPTED BY SOME COURTS OR, IN SOME INSTANCES, FOR EXPERT TESTIMONY

Experts have told the courts, judicial commissions, state legislatures, and the public at large that eyewitness identification is unreliable. It is true that some witnesses identify a filler or no one at all when the culprit is present in an identification procedure. But such inaccuracies do not put an innocent suspect at risk. The real question is: How accurate are eyewitnesses when they identify the suspect whom they have not seen before?

There has been a substantial shift in the field of eyewitness identification since the Special Master issued his report in *Henderson*. Contemporary research strongly indicates that when a witness initially identifies a suspect with high confidence, s/he is highly accurate and hence highly reliable. A witness who initially identifies a suspect with moderate confidence is only slightly less accurate. Moreover, high confidence appears to override most, if not all, of the estimator variables that are said to reduce eyewitness accuracy. Such variables may reduce the proportion of suspect identifications but not the reliability of those that are made, especially with high or moderate confidence.

Experts have testified that there is, at best, a modest relationship between confidence and accuracy. This masked not only the significance of high confidence initial identifications but also low confidence initial identifications. In those Innocence Project exoneration cases where initial confidence was determined, "the witnesses had not been certain at all, a glaring sign that the identification was not reliable." Had the system understood this, the tragic results might have been prevented.

A close reading of the older research (as set forth in the following analyses) shows that it is not as uniform or definitive as it has been presented to be. There are many exceptions and limitations; important findings have been ignored or glossed over; effects may be small and disappear under different circumstances; different studies reach different conclusions; and analytical methods have hidden the information of greatest importance to decision makers, that is, the accuracy of initial suspect identifications.

Because many research findings are being challenged, it would not be appropriate for a court to comment on the eyewitness identifications in the case before it, or to give jury instructions that do not fully and accurately reflect the entire scope of research findings. In essence, jury instructions such as those approved by the New Jersey Supreme Court are tantamount to judicial notice and courts may judicially notice only those facts that are not subject to reasonable dispute. These instructions are subject to reasonable dispute.

The National Academy of Sciences expressed a preference for expert testimony over jury instructions. While endorsing jury instructions as an alternative, it recognized that brief instructions may not provide sufficient guidance but lengthy instructions may be cumbersome and complex. Having said this, it appears that expert testimony on some variables is not reliable and should not be admitted.

The National Academy of Sciences also observed that, in general, jury instructions cause jurors to become more suspicious of all eyewitness identification evidence. A study of the *Henderson* instruction found that they did not improve jurors' ability to discern the quality of the evidence and, instead,

“indiscriminatingly discounted ‘weak’ and ‘strong’ testimony in equal measure.” Altering the results in such a manner does not serve the cause of justice.

RESEARCH DOES NOT PROVIDE A FIRM FOUNDATION FOR JURY INSTRUCTIONS OF THE TYPE ADOPTED BY SOME COURTS AND, IN SOME INSTANCES, FOR EXPERT TESTIMONY

> *I was asked by the prosecution in the Henderson case, which ended up in the New Jersey Supreme Court, and formed the basis for them changing their rules on eyewitness identification. I frankly don't know why . . . he asked me . . . because the other witnesses were friends of mine, were all reading from the same book, and you couldn't get a sheet of paper between what they said and what I said.*

United States v. Thomas, No. 1:13-CR-03897-MV, at 100 (D.N.M. 1/22/15), Testimony of Roy Malpass, Ph.D., the sole social scientist presented by the government in the hearing before the Special Master in *Henderson*. [1]

Introduction

In presentations to the courts, judicial commissions, state legislatures, and the public at large, experts have portrayed eyewitness identification as unreliable. Modern distrust of eyewitness identification accuracy may have started with the Supreme Court's pronouncement that "[t]he identification of strangers is proverbially untrustworthy,"[2] and certainly was fueled/accelerated by the Innocence Project's DNA exonerations (in which mistaken identification was a factor in approximately 70% of the now 349 cases).[3] Into this fertile soil fell research that appeared to show that eyewitness identification is not accurate.[4] To a certain extent, this is true: a witness who identifies a filler or who does not identify anyone when the perpetrator is present is not accurate. But these errors do not increase the risk to an innocent suspect. A witness who identifies an innocent suspect when the perpetrator is not present is not accurate – and it puts the innocent suspect at risk.

Unfortunately, much of the research on which experts have relied failed to distinguish between important and unimportant errors, failed to give proper attention to data that did not accord with the

[1] The defense witnesses in New Jersey were psychology professors Gary Wells, Ph.D., and Steven Penrod, Ph.D., and law professors John Monahan and Jules Epstein. *See* Report of the Special Master, New Jersey v. Henderson (6/18/2010); *State v. Henderson*, 27 A.3d 872, 877 (N.J. 2011). The *Henderson* Instructions were issued by the New Jersey Supreme Court on July 19, 2012, and were effective on September 4, 2012. *See* Supreme Court Releases Eyewitness Identification Criteria for Criminal Cases (July 19, 2012). The Massachusetts Study Group on Eyewitness Evidence met with only a single scientist, Professor Steven Penrod, one of the three researchers who testified in *Henderson*, although it had communications with two others (who were not named). The Report and Recommendations to the Justices Massachusetts Supreme Judicial Court, 2 (July 25, 2014). The Study Group also looked to *Henderson* and *State v. Lawson*, 291 P.3d 673 (Or. 2012), for guidance. *See Commonwealth v. Gomes*, 22 N.E.3d 897 (Ma. 2015).

[2] *United States v. Wade*, 388 U.S. 218, 228 (1967) (quoting then-Professor Felix Frankfurter, *The Case of Sacco and Vanzetti,* The Atlantic (March 1927)). The examples given by the Supreme Court involve unfair lineups (in which the fillers do not resemble the suspect) and other highly suggestive procedures.

[3] The Innocence Project currently reports that 244 cases involved eyewitness misidentification, but apparently this number also includes perjury cases. As a frame of reference, the Innocence Project reports that for cases accepted for DNA testing and closed between 2004-2008, DNA evidence excluded the individual in question in about 43% of cases, included the individual in question in about 42% of cases, and the remaining 15% had no results or were deemed non-probative.

[4] Research has been conducted almost exclusively on the identification of strangers. If a subject knows the target, his/her data is excluded from the study.

stated hypothesis, and failed to apply appropriate methodological approaches and statistical formulas that would inform the police, prosecutors, and courts how likely it is that a *suspect* identification is accurate. It turns out that initial suspect identifications are highly accurate and hence highly reliable.

In the *Henderson* case, the government called only one scientific expert, Dr. Roy Malpass. As the quote above illustrates, his views did not differ from those of the defense experts.[5] Thus, the Special Master in *Henderson* heard only one point of view concerning the research on eyewitness identification – and yet it has had far-reaching effects. There were other researchers at the time who challenged the methodology, the findings, and the applicability of laboratory research to real cases, but they were not heard from.[6] Moreover, a close reading of the research on which the experts relied in *Henderson* and other cases[7] reveals that it is not as uniform or definitive as it has been presented to be. While some factors/variables appear to have validity in some circumstances, there are limitations and exceptions; important findings have been ignored of glossed over; and new research now challenges their underpinnings.

In the years since the Special Master issued his report, there has been a substantial shift in the field of eyewitness identification. Researchers (including those who espoused different views earlier and new ones), using more appropriate methodological approaches[8] and different statistical measures, have concluded that witnesses who initially identify a suspect with high confidence, that is, witnesses who are most likely to testify at trial, are highly accurate.[9] Similarly, witnesses who identify the suspect quickly are highly accurate.

Experts who testified in the past have pointed to a number of factors, called "estimator variables," that, they say, adversely affects accuracy. It is true that these factors appear to reduce overall accuracy. But they do not create false memories for innocent suspects. Instead, it appears that the reduction in accuracy stems primarily from the failure to identify guilty suspects, and not from identifying a greater percentage of innocent suspects. There appear to be fewer identifications when estimator variables are present, but for witnesses who identify the suspect with high confidence, no less accurate ones.

---

[5] Dr. Malpass parted company with his friends only on the issue of simultaneous versus sequential presentation of arrays and the Court found "insufficient authoritative evidence . . . to make a finding in favor of either procedure." *State v. Henderson*, 27 A.3d at 902.

[6] The applicability of eyewitness identification research to real cases (and even that appellation) has been questioned for more than a quarter century. *See* Yuille & G. Wells, *Concerns About the Application of Research Findings: The Issue of Ecological Validity*, 118, 122-123 (1991) ("Caution should be used in generalizing from controlled research studies to real world contexts."); Yuille, *We Must Study Forensic Eyewitnesses to Know About Them*, Am. Psychol., 572, 573 (May 1993) ("[T]here are a handful of studies of forensic witnesses . . . too few to draw any broad conclusions . . . .").

[7] *New Jersey v. Henderson*, 27 A.3d 872 (2011); *State v. Lawson*, 291 P.3d 673 (Or. 2012); *Commonwealth v. Gomes*, 22 N.E.3d 897 (Ma. 2015); *see also State v. Guilbert*, 49 A.3d 705 (Ct. 2012).

[8] *See, e.g.*, National Academy of Sciences, Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 119 (2014) (hereinafter NAS) ("The committee therefore recommends a broad exploration of the effects of different system variables . . . and estimator variables . . . and – importantly – interactions between these variables using either the ROC approach or other tools for evaluation of binary classifiers that can be shown to have advantages over existing analytical methods.")

[9] *See, e.g.*, Wixted & G. Wells, *The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, Psychological Science in the Public Interest, 10, 22 (2017). Even people who identify suspects with a lower level of confidence are much more accurate than they gauge themselves to be.

Witnesses who initially identify the suspect with low confidence also communicate important information to the criminal justice system, that is, that the witness did not get a good enough look and the identification, standing alone, should not support arrest and prosecution, or that the suspect is not the perpetrator. Clearly, in the DNA exoneration cases, an innocent suspect was identified. In every single case in which Professor Brandon Garrett of the University of Virginia was able to obtain sufficient information to ascertain the witnesses' confidence at the time of the initial identification, "the witnesses had not been certain at all, a glaring sign that the identification was not reliable."[10] But the system failed to correctly interpret what the witnesses were saying at the time. Of course, by the time the cases went to trial, the witnesses were highly confident – and by that time they were wrong.

Expert testimony has many advantages over instructions on eyewitness identification.

Through a series of cases starting with *Benn v. United States*, 978 A.2d 1257 (D.C. 2009), the Court of Appeals has removed any barriers to the introduction of expert witnesses on eyewitness identification where the basic criteria of *Motorola* are met.[11] The presentation of an expert witness by the defense affords both the defense and the prosecution the opportunity to explore the strengths and limitations of what is referred to as "eyewitness memory" research. Through direct and cross-examination, a jury is exposed to information that, alone or in combination with other evidence, can be used to assess whether the defendant is the person who committed crime(s) with which s/he is charged. Of course, the jury is free to disregard expert testimony in whole or part, if it so chooses. See D.C. Criminal Jury Instructions for the District of Columbia, No. 2.215 (11th ed. 2005).

The National Academy of Sciences recognized that:

> Expert testimony on eyewitness memory and identifications has many advantages over jury instructions as a method to explain relevant scientific framework evidence to the jury: (1) Expert witnesses can explain scientific research in a more flexible manner, by presenting only the relevant research to the jury; (2) Expert witnesses are familiar with the research and can describe it in detail; (3) Expert witnesses can convey the state of the research at the time of the trial; (4) Expert witnesses can be cross-examined by the other side; and (5) Expert witnesses can more clearly describe the limitations of the research.[12]

The NAS endorsed jury instructions as an *alternative* when expert testimony was not available. However, it recognized that "Brief instructions may not . . . provide sufficient guidance . . . but lengthy instructions may be cumbersome and complex." NAS at 43.

---

[10] Garrett, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG, 64 (Harvard U. Press 2011) (In 40% of the trials, the witnesses did not initially identify the defendant; in 21%, the witness admitted to initial uncertainty; in 9%, the witnesses reported not having seen the culprit's face at all. There were multiple reasons in some cases).

[11] Benn referred specifically to *Dyas v. United States*, 376 A.2d 827 (D.C. 1977), and *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923), but that standard was replaced by *Motorola v. Murray*, 147 A.3d 751 (D.C. 2016).

[12] National Academy of Sciences, Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 40 (2014) (NAS) (citations omitted).

Courts should not comment on the evidence or take judicial notice of contested facts

The defendant is not entitled to have a jury instruction that presents only the defense position on the accuracy of eyewitness identification. The proposed instructions are unbalanced, capture only a portion of the research, do not take into account the limitations and the external validity of the laboratory literature itself, do not address contrary findings both within the academic literature and in real-world studies, and do not reflect the most recent research. Jury instructions place the court in the position of being an expert witness without the opportunity to more fully explore and challenge the bases of its opinions.

It may be that the court could find sufficient evidence of reliability for an expert to testify about various issues that are said to affect eyewitness identification.[13] Such a finding, however, would support only the admissibility of the expert's opinion; it would not establish that it was uncontroverted or indisputable or irrefutable. It is one thing to permit jurors to hear from an expert on the psychology of eyewitness identification subject to cross-examination and argument. It is quite another to have the Court give instructions that incorporate one view of the research.

Unlike expert testimony that the jury can accept or reject, *see* Jury Instruction 2.215 ("You are not bound by an expert's opinion."), the jury must follow the judge's instructions, *see* Jury Instruction 2.101 ("It is your duty to accept the law as I instruct you. . . . You may not ignore or refuse to follow any of [the instructions]"). As the Court of Appeals has cautioned, "[t]he judge's emphatic instruction . . . is likely to [be] taken by the jury as a legal injunction ('the law as I state it to you'), . . . an injunction that the jury had 'a duty to accept' without 'question[ing its] wisdom,' rather than as referring to the 'facts' as to which the jurors were the sole and exclusive judges.'" *Wheeler v. United States,* 930 A.2d 232, 244-245 (D.C. 2007); *see Boyde v. California*, 494 U.S. 370, 384 (1961) ("judicial instructions are likely to be given greater weight than advocacy because jurors perceive instructions "as definitive and binding statements of the law.").

It would be difficult for the jury to understand what the law is and what the facts are when they both come out of the court's mouth. Moreover, "[i]t is primarily the task of counsel, not the court, to develop facts essential to jurors understanding of the case." *Davis v. United States*, 567 A.2d 36, 39 (D.C. 1989). For a judge to present facts, particularly contested facts, in the guise of "scientific research," would constitute an impermissible comment on the evidence. *See Wheeler v. United States, supra*, 930 A.2d at 243–44 (explaining that the trial judge's common-law privilege to comment on the evidence "has inherent limitations and must be exercised cautiously, for a judge's influence on the jury is necessarily and properly of great weight and his or her lightest word or intimation is received with deference, and may prove to be controlling") (internal quotation marks, brackets, and citations omitted), *cited in Blaine v. United States*, 18 A.3d 766, 785 (D.C. 2011).

[13] "'When a trial court, applying [Rule 702], rules that an expert's testimony is reliable, this does not necessarily mean that contradictory expert testimony is unreliable. [Rule 702] is broad enough to permit testimony that is the product of competing principles or methods in the same field of expertise.' Fed. R. Evid. 702 advisory committee's notes to 2000 amendments. Indeed, we expect that many cases will feature expert witnesses espousing different views of the evidence. Their testimony will be tested by the adversary process and evaluated by the jury." *Motorola Inc. v. Murray*, 147 A.3d 751, 757 (D.C. 2016).

If the court is going to comment on the evidence (or lack of evidence), then clearly it would have an obligation to do so fully and accurately, informing the jury of all of the countervailing considerations that might lead to different conclusions. The proposed instructions – and others like them – fail on all counts. Ultimately, instructions that adopt a particular view of the research are the equivalent of judicial notice, and "[t]he court may judicially notice [only] a fact that is not subject to reasonable dispute . . . ." Fed R. Evidence 201; *Gee v. United States*, 54 A.3d 1249, 1266 (D.C. 2012) ("A trial court errs if it takes judicial notice 'without determining that ... the sources relied upon have an accuracy that cannot reasonably be questioned[.]'") (internal citations omitted). Here the scientific research relied upon has an accuracy that can reasonably be questioned, and we do so.

The current instructions on identification and witness credibility tell the jury the kinds of information it may consider in assessing witness testimony and particularly eyewitness testimony. *See* Jury Instructions 9.210 and 2.200. The current instructions wisely do not tell jurors how to consider this information. Going beyond these instructions risks having the court weigh in on one side of a contested issue by providing incomplete and biased information that may or may not apply to the witnesses in the case, or to any witnesses at all.[14]

Jury Instructions based on flawed or outdated research should not be given

In general, jury instructions on disputed eyewitness research should not be given. In particular, the instructions that have been proposed or adopted in other jurisdictions are seriously flawed. The following are some general comments on the instructions:[15]

(1) As the National Academy of Sciences observed, "[R]esearch findings on the effectiveness of jury instructions on assessment of eyewitness identification evidence have been mixed. In general, such studies find that jury instructions cause jurors to become more suspicious of all eyewitness identification evidence. A recent study of the effect of the New Jersey jury instructions used in *Henderson* found that the instructions reduced reliance on *both* strong and weak eyewitness identification evidence."[16] In other words, the *Henderson* instruction did not improve jurors' ability to discern the quality of an identification; instead jurors "indiscriminatingly discounted 'weak' and 'strong' testimony in equal measure."[17] Impugning an entire category of evidence in such a fashion is not fair to the victims and witnesses, the government, or the cause of justice.

(2) While we recognize the defense has the right and the duty to challenge each piece of evidence used to establish identity, it would be inappropriate for the Court to suggest that the jury should treat

[14] Given some of the research discussed here and in other sections, it may be prudent to reconsider the current instructions. For example, if there is little or no difference in high confidence identifications for same and cross-race individuals, race may not be something that should be considered.

[15] Analyses of the research on specific variables follow: confidence and accuracy, cross-race, stress, weapon focus, distance, lighting, exposure duration, retention interval, disguise, and intoxication.

[16] NAS at 43 (citations omitted).

[17] Papailiou, Yokum & Robertson, *The Novel New Jersey Eyewitness Instruction Induces Skepticism But Not Sensitivity*, Social Sciences Research Network (August 5, 2014), published on line, PLOS (12/9/2015). See also Dillon, Jones, Bergold, Hui & Penrod, *Henderson Instructions: Do They Enhance Evidence Evaluation?*, J. Forensic Psych. Research & Practice 1, 12 (published on line 1/13/2017) ("[R]ather than increasing jurors' sensitivity to witnessing and identification conditions, the *Henderson* instructions induced skepticism by reducing convictions regardless of eyewitness quality.").

eyewitnesses skeptically, or that identification evidence should be scrutinized more carefully than other evidence. Research does not support such a conclusion with respect to honest witnesses who identify a suspect with high confidence (or even lesser confidence) initially. Expert witnesses, cross-examination, and argument are the appropriate means by which to explore the strengths and weaknesses of a given witness's ability to identify the perpetrator.

(3) The proposed instructions generally provide blanket statements that capture neither the variability of the research nor its limitations. Instructing the jury properly about "research" is particularly fraught with peril when the effects found in the research are small and/or either do not exist or disappear under different circumstances; when researchers themselves say additional research is necessary; when some effects are trumped by others; when the effects move in opposite directions; when researchers relied on protocols that are not used by law enforcement; when research produces inconsistent results; when research that does not find the effect at issue is rejected for publication; when old research does not support the proposition for which it is cited; and when new research methods and analytical approaches raise significant questions about past findings.[18] Differences between the conduct of laboratory research and actual cases further complicate matters.

(4) Recommendations in *Henderson* and other cases were based in part on a "consensus" contained in a 2001 survey to which 64 experts responded.[19] *Henderson*, 27 A.3d 911-912. As one article commented, "all these surveys prove is that those who believe in laboratory-based research as a valid context to study eyewitness memory share some common beliefs."[20] As the accompanying analyses of specific issues illustrate, a consensus among knowledgeable and objective researchers likely would be different.

(5) A careful review of the studies, particularly older studies, discloses that:

- Some studies have not followed the "best practices" recommended for the police (such as, the target "may or may not" be in the array or offering a "don't know" or "not sure" option).

- Many studies have not weeded out subjects who (if asked) might say, "I didn't get a good enough look at him," or "I was looking at the gun the whole time," or "I wasn't paying attention to his face," responses that would lead the police away from conducting an identification procedure.

- In some videotape or mock event (not necessarily mock crime) studies, it is not clear how many subjects actually saw or paid attention to the target and, if so, for how long.

[18] For example, the much touted claim of a "sequential advantage" in the display of photo arrays has been significantly undermined and the National Academy of Sciences declined to endorse the exclusive use of this method, s*ee* NAS at 104, as did the Department of Justice in its *Memorandum on Eyewitness Identification: Procedures for Conducting Photo Arrays* (January 6, 2017).

[19] Kassin, Tubb, Hosch & Memon, *On the ''General Acceptance'' of Eyewitness Testimony Research: A New Survey of the Experts*, Am. Psychol., 405, 407 (2001).

[20] Yuille & Cooper, *Challenging the Eyewitness Expert*, in Ziskin & Faust (Eds.), COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY, 685, 695, (6th edition 2012). *See* Yuille, Daylen, Porter & Marxsen, *Challenging the eyewitness expert*, In J. Ziskin (Eds.), COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY, 1266, 1289 (1995) ("depending on the research paradigm employed, widely varying conclusions can be drawn about specific eyewitness factors.").

- In facial recognition studies, subjects view multiple photographs for milliseconds to a few seconds each.

- Some studies use dozens of photographs or videos of targets whom the subjects are later to identify, quite unlike most crimes that involve a more limited number of perpetrators.

- Some arrays in which the real target is not present use a substitute who closely resembles the real target, something the police cannot do deliberately because they do not know what the perpetrator looks like. It is only by chance that an innocent suspect would closely resemble, but not be, the perpetrator.

- Laboratory research does not (because it cannot) create the same emotional response as real crimes, as is demonstrated by field studies.

- Some studies use photographs for the initial exposure although live human beings certainly would be viewed at the time of a crime.

- Laboratory subjects are aware that no consequences are attached to their choice or failure to choose.

- Until recently, laboratory research subjects were not necessarily asked about their level of confidence at the time of the identification; even when confidence has been solicited, researchers often report only average accuracy rates, which say little, if anything, about the accuracy of witnesses who select a target or suspect at different levels of confidence.

- Researchers do not necessarily provide (or keep) all of the data underlying their conclusions.

- Researchers sometimes minimize results that are inconsistent with their hypotheses or claim statistical significance when the results are not practically significant or statistically trustworthy.

- Scientific journals do not necessarily publish results that are inconsistent with other findings or that find no effect, thereby skewing the significance of published studies.

- Many studies are small and/or have not been replicated. Some studies that have been replicated reach opposite conclusions.

(6) If current research were appropriately considered, jury instructions would include, for example: (a) "Confidence at the time of the initial identification is a reliable indicator of accuracy; if anything, witnesses are likely to be under-confident rather than over-confident;" (b) "stress does not increase false identification;" or (c) "delay between the crime and the identification does not undermine the reliability of the identification."

(7) The *Manson v. Brathwaite* factors have been criticized as not being "diagnostic of reliability."[21] But at least "*the level of certainty demonstrated at the confrontation*," turns out to be correct. 432 U.S. 98, 114 (1977). In context, the word "confrontation" means the initial identification, not an in-court identification. *See id*. at 106 (using the word "confrontation" to mean the identification procedure); *id*.

[21] NAS at 6 (citing critics).

at 108, 110 (discussing "identifications resulting from 'unnecessarily suggestive confrontation'"); *id*. at 115 (only two days elapsed between the crime and the "confrontation"). The phrase "the level of certainty demonstrated at the confrontation" seems to brook only one interpretation, that is, the level of certainty demonstrated or expressed at the initial identification. Whether or not this is a correct interpretation of *Manson*, the level of certainty demonstrated at the time of the initial identification, if honestly given, is likely to be more informative than a changed level of certainty given later.[22]

In the sections that follow, we present an overview of the old and new research on system and estimator variables, followed by more detailed analyses of the research on confidence and accuracy, stress, weapon focus, exposure duration, distance and lighting, retention interval (delay), disguise, and intoxication. A glossary is included at the end.

[22] *See* NAS at 111 (judges should "make juries aware of all prior identification procedures").

AN OVERVIEW OF THE RESEARCH ON EYEWITNESS IDENTIFICATION OF STRANGERS, NEW AND OLD[23]

Important Points

(1) The record in *Henderson* was not developed enough for the Special Master to make truly informed recommendations. *Henderson* also commenced seven years ago, and much has happened in the meantime.
(2) For cases in which an eyewitness has identified the suspect, the criminal justice system needs to know how reliable a *suspect* identification is at various levels of confidence. Research that includes fillers (and non-IDs) and does not segregate accuracy by confidence levels does not tell juries what they need to know.
(3) A highly confident suspect identification at the time of the initial identification (at least in a fair lineup, with no influence, by an honest witness) is highly reliable (typically 95-100%) and a moderately confident suspect identification is only slightly less so (typically 90-95%).
(4) Virtually all of the recent research indicates that the reliability of a high-confidence initial suspect identification is not adversely affected by estimator variables, that is, the presence of a variable may reduce overall accuracy, but it does not diminish the reliability of high-confidence *suspect* identifications (*e.g.*, fewer suspect IDS, not less reliable ones).
(5) Some variables (*e.g.*, weapon focus, stress) reduce true identifications but apparently do not increase false identifications and, therefore, the presence of such a variable would not put an innocent suspect at a higher risk of being misidentified as the culprit.
(6) The effect of some variables (*e.g.*, cross-race effect) is very small and disappears when, for example, the exposure duration is longer or the witnesses expresses high confidence and, therefore, the presence of such a variable would not put an innocent suspect at a higher risk of being identified as the culprit.
(7) There is no way to tell which variables affect which witnesses.

System Variables

Simultaneous versus sequential presentation

The National Academy of Sciences asked: "[C]an we draw definitive conclusions about which lineup procedure (sequential or simultaneous) is preferable? At this point, the answer is no."[24]

A debate has continued with Wells, Steblay, and Dysart (there is a sequential advantage) on one side and Amendola and Wixted (more good IDs, fewer bad ones with simultaneous)[25] on the

[23] Note that all of this applies only to stranger identifications. Although some defense experts are now trying to apply the stranger research to known perpetrators, there is no scientific basis for it. Researchers have not studied the effect of estimator variables on the identification of known perpetrators because it would not be a worthwhile endeavor. The brain encodes, stores, and retrieves known and strange faces differently.

[24] National Academy of Sciences, Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 118 (2014) (NAS). The International Association of Chiefs of Police recently dropped its long-standing recommendation that only sequential procedures should be used. *See* IACP, *Model Policy Concepts & Issues Paper*, 1 (September 2016) ("the sequential and simultaneous approaches are both valid methods of conducting an identification procedure and [this policy] does not recommend one over the other.").

[25] Amendola & Wixted, *The Role of Site Variance in the American Judicature Society Field Study Comparing Simultaneous and Sequential Lineups*, J. Quant. Criminol., 1 (2017); Amendola & Wixted, *Comparing the diagnostic accuracy of suspect identifications made by actual eyewitnesses from simultaneous and sequential lineups in a*

other. Before the NAS Report was issued, other researchers found no evidence for a sequential advantage.[26] Since then, other researchers have found a simultaneous advantage.[27]

Recently, the Department of Justice issued guidance on the use of photo lineup procedures by federal law enforcement agencies, which noted that "there has been an evolution in views on whether the 'sequential' administration of a photo array (presenting the witness one photo at a time) results in more accurate identifications than a 'simultaneous' administration (presenting all of the photos at once)."[28] It recommended the use of either simultaneous or sequential photo lineups while acknowledging that, although earlier research appeared to favor the sequential procedure, more recent research "reached different conclusions, suggesting that simultaneous procedures may result in more true identifications and fewer false ones."[29]

Double blind, blind, blinded administration (unless impracticable)

Based on research in other fields, the NAS recommended double-blind or blinded testing. However, "[t]here remains relatively little evidence evaluating the merits of double-blind lineup administration. Consequently, its status as a reform has more to do with the historical importance of blind testing in other fields than the existence of a definitive empirical base involving lineup testing."[30]

---

randomized field trial., 1, J. Exp. Criminol., 263 (2015); Amendola & Wixted, *No possibility of a selection bias, but direct evidence of a simultaneous superiority effect: a reply to Wells et al*., J. Exp. Criminol., 291 (2015); Wells, Steblay & Dysart, *A test of the simultaneous versus sequential lineup methods: an initial report of the AJS national eyewitness identification field studies (*American Judicature Society 2011); Wells, Steblay & Dysart, *Double-blind photo-lineups using actual eyewitnesses: an experimental test of a sequential versus simultaneous lineup procedure*. Law & Hum. Behav., 1 (2015); Wells, Steblay & Dysart, *The flaw in Amendola and Wixted's conclusion on simultaneous versus sequential lineups,* J. Exp. Criminol., 285 (2015). *See also* Wixted, Mickes, Dunn, Clark & W. Wells, *Estimating the reliability of eyewitness identifications from police lineups,* Proceedings of the National Academy of Sciences, 304, 304 (2016).

[26] Gronlund, Carlson, Dailey & Goodsell, *Robustness of the sequential lineup advantage.* J. Exp. Psychol.: Appl., 140, 140 (2009) ("there appears to be no strong preference for conducting lineups in either a sequential or a simultaneous manner"); Malpass, Tredoux & McQuiston-Surrett, *Public policy and sequential lineups*, Legal & Criminol. Psychol., 1, 1 (2009) ("the corpus of research on sequential lineups does not satisfy the needs of policy sufficiently to justify its mandated use as the required identification procedure throughout the criminal justice system."); Carlson, Gronlund & Clark, *Lineup Composition, Suspect Position, and the Sequential Lineup Advantage*, J. Exp. Psychol., 118, 118, 126 (2008) ("A sequential lineup advantage was found only for unfair lineups."). *See also* Mickes, Flowe & Wixted, *Receiver Operating Characteristic Analysis of Eyewitness Memory: Comparing the Diagnostic Accuracy of Simultaneous Versus Sequential Lineups*, J. Exp. Psychol.: Appl., 361, 375 (2012) (finding evidence for a simultaneous superiority effect);

[27] Dobolyi & Dodson, *Eyewitness Confidence in Simultaneous and Sequential Lineups: A Criterion Shift Account for Sequential Mistaken Identification Overconfidence*, J. Exp. Psychol.: Appl., 345, 355 (2013) ("Overall, we show that sequential lineups are both less accurate and produce higher confidence false identifications than do simultaneous lineups.").

[28] Deputy Attorney General Sally Q. Yates, *Eyewitness Identification Procedures for Conducting Photo Arrays*, 1 (U.S. Department of Justice, January 6, 2017).

[29] Yates (2017), *supra*, at 6-7.

[30] Gronlund, Mickes, Wixted & Clark, *Conducting an eyewitness lineup: How the research got it wrong*, in THE PSYCHOLOGY OF LEARNING AND MOTIVATION, 1, 30 (Ross, ed. 2015). *See also* Greathouse & Kovera, *Instruction Bias and Lineup Presentation Moderate the Effects of Administrator Knowledge on Eyewitness Identification,* Law & Hum. Behav., 70, 81 (2009) ("This study illustrates that there are still many questions about the effects of administrator

Hence, adverse consequences (in the form of exclusion, limitation, or jury instructions) should not be attached to the failure to administer photo arrays in a double-blind, blind, or blinded manner, especially where it was impracticable.

There is research that indicates that witnesses can be explicitly induced to pick a particular photograph and/or that feedback before a confidence statement is taken can inflate confidence levels.[31] It is easier to establish that neither of these circumstances occurs with blind administration; video recording the procedure also could be used to the same effect. The same results can be had, however, by not saying anything to the witness after s/he has made an identification.[32] Note that deliberately trying to influence a witness reduces the witness's confidence and, therefore, limits the witness's use as an identification witness at trial.[33]

Witness instructions

Some experts claim that the failure to instruct a witness that the perpetrator "may or may not" be in the array could affect the reliability of an identification. This is based on studies comparing "biased" and "unbiased" instructions. These terms have been defined somewhat differently by different researchers. One set says: "Biased lineup instructions are those that either fail to warn the witness that the culprit might not be in the lineup or imply that the culprit is in the lineup. Unbiased instructions, in contrast, warn the witness that the culprit might not be in the lineup."[34] The "may or may not" instruction is considered unbiased.

---

knowledge of a suspect's identity and double-blind lineup administration on witness behavior that remain unanswered before solid policy recommendations can be made."). The little research that exists does not indicate that blind administration reduces false identifications. Phillips, McAuliff, Kovera & Cutler, *Double-Blind Photoarray Administration as a Safeguard Against Investigator Bias*, J. Appl. Psychol., 940, 947 (1999) ("Participant administrator's knowledge of the suspect's identity had a biasing effect in sequential photo arrays only . . . and then only when an observer was present. For observer-absent photoarrays, there were no differences in false identification rates for the single-versus double-blind procedures, irrespective of whether they were presented sequentially or simultaneously."); Perlini & Silvaggio, *Eyewitness Misidentification: Single vs Double-Blind Comparison of Photospread Administration*, Psychol. Reports, 247, 252, 253 (2007) ("[t]here was no significant effect for photospread procedure (blind, informed) . . .").

[31] *See e.g.,* G. Wells & Bradfield, *"Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing* experience, J. Appl. Psychol., 360 (1998); Dysart, Lawson & Raine, *Blind Lineup Administration as a Prophylactic Against the Postidentification Feedback Effect*, Law & Hum. Behav., 312 (2011). There appears to be only one study on inadvertent influence. Garrioch & Brimacombe, *Lineup administrators' expectations: Their impact on eyewitness confidence*, Law & Hum. Behav., 299 (2001). Because the number of subjects in each condition was small, and the standard deviation estimates vary widely, the results are questionable. Moreover, they have not been replicated. Thus, this study is a weak reed upon which to depend for any conclusions about inadvertent or subtle influence.

[32] Instructions like, "Do not assume that I know who the suspect is," and "I cannot tell you if you have picked the right person or the wrong person," also could alert the witness not to look to the administrator for confirmation.

[33] Clark, Brower, Rosenthal, Hicks & Moreland, *Lineup administrator influences on eyewitness identification and eyewitness confidence*, J. Appl. Res. Mem. & Cog., 58 (2013) (deliberately steering a subject to the suspect resulted in more identifications, but false suspect identifications were made with lower confidence). In an earlier study, using a scale of 1-10, the mean confidencefor the non-blind feedback condition was 7.53, below what would be considered high confidence. Dysart et al. (2011), *supra*, at 317, Table 1.

[34] Wells, Yang, Smalarz, *Eyewitness Identification: Bayesian Information Gain, Base-Rate Effect-Equivalency Curves, and Reasonable Suspicion*, Law & Hum. Behav., 99, 109 (2015).

Most of the studies compare the unbiased "may or may not" instruction with a biased instruction that states or implies that the culprit is in the array.[35] However, one study compared the unbiased instruction with the "neutral" instruction: "If you see the person from the video in the lineup, please pick him; otherwise, choose the 'not present' option."[36] It found no difference between the unbiased ("may or may not") instruction and the "neutral" instructions. Another study compared the unbiased instruction with a "minimal" instruction: "In a moment I am going to show you a group of photographs. When you have looked at all the photos, indicates below whether or not you see the person who committed the crime. . . ."[37] It found that "numerically, the minimal instruction produced both the highest hit rate when the offender was present ([47] percent), when a choice was made and the lowest false alarm rate when the offender was absent (30 percent)."[38] The authors opined, "[p]erhaps eyewitnesses, left to their own inherent judgment, are more competent decision makers than they are sometimes given credit."[39] These researchers also noted, "in order to obtain a significantly higher rate of false identifications when the offender was absent from the photospread, subjects had to be blatantly misled that the offender was in fact present in the photospread."

Two studies have found that "Biased instructions increased confidence unless the [perpetrator] was absent and lineup members were similar, where they decreased confidence."[40]

The "police will continue to investigate" instruction has not been studied, but it seems unobjectionable. This and the "may or may not" are the only two instructions mentioned by the NAS. Two other instructions recommended by the defense (that also have not been studied) are objectionable. First, it is not the job of an eyewitness to "clear the innocent or implicate the guilty." That is the job of the police, prosecutors, and the court based on all of the evidence in the case, not on the identification—or lack of identification – by one eyewitness. Second, a witness should identify the perpetrator if he can. Fear of reprisal, fear of being labelled as a "snitch," ties to perpetrators or their families or friends, or simply a reluctance to get involved, mean that many witnesses do not come forward and/or are reluctant to cooperate. Encouraging those who have been identified as witnesses to "opt out" by instructing them that they are not "required to make an identification" does not advance the cause of justice or the protection of the community. Although the effect of such an instruction has not been scientifically investigated, it is difficult to imagine how it would increase the accuracy of eyewitness identifications. If it has any effect at all, it most likely would discourage identifications from being made, whether the suspect is innocent or guilty.

---

[35] Malpass & Divine, *Eyewitness Identification: Lineup Instructions and the Absence of the Offender*, J. of Applied Psych. 482 (1989) ("We believe that [the culprit] is present in the lineup . . . which of these is the person you saw" committing the crime."); Leippe, Eisenstadt & Fauch, *Cueing Confidence in Eyewitness Identifications: Influence of Biased Lineup Instructions and Pre-Identification Memory Feedback Under Varying Lineup Conditions*, Law Hum. Behav. 194, 198 (2009) ("look carefully . . . and determine which person is the thief . . .and . . . click on your selection) (elipses in original).

[36] Mickes, Seale-Carlisle, Wetmore, Gronlund, Clark, Carlson, Goodsell, Weatherford & Wixted, *ROCs in Eyewitness Identification: Instructions vs. Confidence Ratings*, XX (2017)

[37] Paley & Geiselman, *Effects of photospread Instructions*, Am. J. Forensic Psych, 3, 7 (1989).

[38] *Id*. at 12. The text has 74% but Table 1 has 47%. Table 1 appears to be correct. Either way, the percentage of correct hits was greater, although only slightly greater, with "minimal" instructions than two "may or may not" instructions.

[39] *Id.* at 12.

[40] Leippe et al. (2009), *supra* at 194.

Estimator Variables

Confidence and accuracy

Latest research

Confidence is strongly related to accuracy. "[T]he idea that initial confidence is not strongly related to accuracy conflicts with virtually all empirical evidence – both in the lab and in the real world – that has accumulated over the last 15 to 20 years. . . . These studies have established beyond any reasonable doubt that, for adults who make an ID from a lineup, the relationship between initial confidence and accuracy in a typical forensically-relevant lab study – *precisely the kind of study that once convinced the field that the relationship is weak* – is in fact strong."[41]

How reliable are suspect IDs? The question before the court in a trial is not how accurate eyewitnesses are overall. The question is how accurate, and hence how reliable, witnesses are who identify *suspects* with high confidence. Suspects who are not chosen are not prosecuted (at least not on the basis of eyewitness identification), nor are fillers. Suspects identified with less than high confidence are not prosecuted unless there is other evidence of identification (DNA, fingerprints, GPS, surveillance cameras, tag readers, possession of stolen property, etc.).

New research on confidence and accuracy appropriately excludes fillers since they are known to be innocent. Using a statistical method (the confidence-accuracy characteristic (or CAC) that does not include fillers – who are not prosecuted – leads to the conclusion that highly confident witnesses are highly accurate.[42] *See* Glossary (CAC).

Initial confidence statements are significant. Defense experts have said that obtaining a confidence statement at the time of the initial identification is necessary, but they then obscured the significance of confidence statements – both high and low -- by using outdated statistical methods that masked their significance. [43] *See* Glossary (point bi-serial correlation coefficient).

Older research

---

[41] Wixted, Mickes, Clark, Gronlund & Roediger III, *Initial Eyewitness Confidence Reliably Predicts Identification Accuracy*, American Psychological Assn., 515, 516, 518 (September 2015) (emphasis in original). *See also* Dobolyi & Dodson, *Eyewitness Confidence in Simultaneous and Sequential Lineups: A Criterion Shift Account for Sequential Mistaken Identification Overconfidence,* J. Exp. Psychol.: Appl., 345, 345 (2013) (citations and parentheticals omitted) (there is a "growing number of studies [showing that] the magnitude of an individual's confidence rating in a lineup decision can be well calibrated with its likely accuracy.").

[42] Wixted & G. Wells, *The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, Psychological Science in the Public Interest, 10, 24 (2017); Mickes, *Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness mem*ory, J. Appl. Res. Mem. & Cog., 93, 101 (2015); Wixted, Mickes, Dunn, Clark & W. Wells, *Estimating the reliability of eyewitness identifications from police lineups,* Proceedings of the National Academy of Sciences, 304, 304 (2016).

[43] High confidence is marked by 80%-100% or 90-100%s, and statements like "That's him!," "positive," "definitely him," "I'll never forget that face," and "I'm sure." Low confidence would be 60% or less and words like "looks similar," "possibly the guy," "maybe," "I think but I am not sure," and "not too sure." Moderate confidence is 70-80% and words like "very similar," "looks like," "looks most like," "pretty sure," "I think it's him," and "pretty certain."

A number of older studies concluded there was a strong relationship between confidence and accuracy.[44]

Speed of Identification

Latest research

Fast = accurate. "Fast identifications (6 seconds or less) and confident (90-100%) individuals showed an impressive 97% accuracy rate when they selected someone from a lineup."[45]

Older research

Fast = accurate. Older research has come to similar conclusions with somewhat longer time frames and somewhat lower accuracy rates[46] – although the researchers used statistical methods that did not focus solely on suspect identifications and therefore underestimated the accuracy rate. There appears to be no dispute that "jump-out identifications" are "desirable witness expressions of absolute certainty."[47]

Stress

Latest research

Anxiety does not affect false alarm rates. Two studies indicate that stress decreases true identifications, but does not affect false identifications. In one, the "results indicate that anxiety

[44] Behrman & Richards, *Suspect/Foil Identification in Actual Crimes and in the Laboratory: A Reality Monitoring Analysis*, Law & Hum. Behav., 279, 297 (2005) (a mixed archival and laboratory study concluded that "witnesses who display high levels of certainty . . . are unlikely to choose innocent persons."); Palmer, Brewer, Weber & Nagesh, *The Confidence-Accuracy Relationship for Eyewitness Identification Decisions: Effects of Exposure Duration, Retention Interval and Divided Attention*, American Psychol. Assn., 55, 56 (2013) ("[T]here is a growing body of evidence that points to a meaningful CA relationship under certain conditions. For example, confidence has been shown to be a useful predictor of accuracy when . . . the witness makes a positive identification as opposed to rejecting the lineup."); Lindsay, Read & Sharma, *Accuracy and Confidence in Person Identification, the Relationship is Strong When Witnessing Conditions Vary Widely,* Am. Psychol. Soc., 215, 217 (May 1998) ("These considerations [variability in exposure duration, retention interval, unusual appearance, difference in appearance] suggest that ability to identify the perpetrator varies greatly from one real-world witness to another, and thus, in turn, suggests that the real-world AC relationship is strong.").

[45] Dodson & Dobolyi, *Confidence and eyewitness identifications: The cross-race effect, decision time and accuracy,* Appl. Cog. Psychol., 113, (2016), *citing* Sauerland & Sporer, *Fast and confident: postdicting eyewitness identification in a field study*, J. Exp. Psychol.: Appl., 646 (2009).

[46] Weber, Brewer, Wells, Semmler & Keast, *Eyewitness Identification Accuracy and Response Latency: The Unruly 10-12-Second Rule*, Am. Psychol. Assn., 139, 146 (2004) ("the combined use of 90-100% confidence and the 10-[second] time boundary diagnosed identification decisions with a high probability of accuracy (88.1%) overall."); Dunning & Perretta, *Automaticity and eyewitness accuracy: a 10-to-12 second rule for distinguishing accurate from inaccurate positive identifications*, J. Appl. Psychol., 951, 951 (2002). (Identifications made within 10-12 seconds were 90% accurate compared to 50% for those that took more time.); Pickering & Darling, *Characteristics of Eyewitness Identification that Predict the Outcome of Real Lineups,* Appl. Cog. Psychol., 969, 984 (2003) ("Fast decisions are more likely to result in identification of the suspect (87%) than average or slow decisions (38% and 31% respectively.").

[47] Klobuchar, Steblay & Caligiuri, *Improving Eyewitness Identifications: Hennepin County's Blind Sequential Lineup Pilot Project*, Cardozo Public Law, Policy & Ethics J., 381, 399-400 (April 2006).

degrades performance in a face-matching task, but only with respect to hits, not false alarms. This finding . . . provides further support for the dissociation between the ability to accurately identify a genuine face match (i.e., anxiety lowers hit rates) and the ability to accurately identify a lack of a match (i.e., anxiety does not affect false alarm rates)."[48] In the other, "emotional participants were more likely to incorrectly identify [a filler]" in a target present lineup but "there was no association between the experience of emotion and lineup decision" in a target absent lineup.[49]

"Stress had no impact on identification performance in target-present or target-absent line-ups."[50]

Older research

Stress does not increase false identifications. One of the two studies primarily relied on by the Special Master in *Henderson* stated, "[t]he overall negative impact of heightened stress on accuracy of face identification was *due entirely* to a substantial effect on hit rate for TP [target-present] lineups. The correct rejection rate for TA [target-absent] lineups was unaffected by stress level."[51] The other study, frequently characterized as being more ecologically valid, actually showed that there were slightly more false identifications in the low stress condition than in the high stress.[52] This means that an innocent suspect is NOT more likely to be identified in a high stress condition.

Stress effects mostly attributable to one study. In the major meta-analysis on stress, there was virtually no difference in the proportion of correct IDs in the high (.56) and low (.58) stress conditions in facial recognition studies. To the extent that "more ecologically valid" eyewitness ID studies were examined, one study, which utilized a live staged crime, "was responsible for most of the difference in effect sizes."[53] Another study (not in the meta-analysis), which utilized a similar live staged crime found that "to the extent that there was a correlation [between stress

[48] Attwood, Penton-Voak, Burton & Munafo, *Acute Anxiety Impairs Accuracy in Identifying Photographed Faces*, Psychol. Science, 1591, 1593 (2013).

[49] Houston, Clifford, Phillips & Memon, *The Emotional Eyewitness: The Effects of Emotion on Specific Aspects of Eyewitness Recall and Recognition Performance*, Am. Psychol. Assn., 118, 124 (2013).

[50] Sauerland, Raymaekers, Otgaar, Memon, Waltjen, Nivo, Slegers, Broers & Smeets, *Stress, stress-induced cortisol responses, and eyewitness identification performance*, Behav. Sci. Law, 580, 590 (2016).

[51] Deffenbacher, Bornstein, Penrod & McGorty, *A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory,* Law & Hum. Behav., 687, 695 (2004) (emphasis added).

[52] Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Baranoski & Southwick, *Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress*, Int'l J. Law & Psychiatry, 265 (2004). For TP lineups, high stress reduced ID accuracy (failing to pick the true interrogator) for 46% of the subjects, had no effect on 43% of the subjects and increased accuracy for 11% of the subjects. For TA lineups, there were more false identifications (failing to reject the array) in the low stress condition (46%) than in the high stress condition (39%). These percentages were not disclosed or discussed in the article but can be determined from the data presented. *See id*. at 272, Table 1 (false alarms are the inverse of true negative responses). *See also* Clark & Wells, *On the Diagnosticity of Multiple-Witness Identifications*, Law & Hum. Behav., 406, 415 (2008) ("Morgan showed that stress reduced the correct identification rate in target present lineups, but had no effect on the mistaken identification rate in TA [target absent] lineups.").

[53] Deffenbacher, Bornstein, Penrod & McGorty, *A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory,* Law & Hum. Behav., 687, 695 (2004), *citing* Buckhout, Alper, Chern, Silverberg & Slomovits, *Determinants of eyewitness performance on a lineup*, Bull. Psychon. Soc., 191, 191 (1974).

and ID accuracy], it was in the direction of higher arousal [*e.g.*, greater stress] being associated with more accurate identification."[54]

Laboratory studies cannot mimic real crimes and, therefore, laboratory stressors are often external to the event. Because they cannot ethically replicate real crimes,[55] laboratory researchers often use "extraneous sources of arousal like loud noise, failure stress, worry about a threatening experimental situation, and so on . . . ." that are likely to "distract him or her" from the event at issue,[56] and thus reduce ID accuracy.

Stress affects different people differently. One study, using a single stressor, divided the subjects down the middle: the half that was above the median ("high anxiety") made fewer accurate identifications (29%) than the half that was below the median ("low anxiety") (81%). This study demonstrates that one cannot assume that a certain situation would necessarily generate high stress.[57]

High confidence = high accuracy. For participants in this study who made an identification -- whether high or low stress – the "accuracy of identification was reliably associated with confidence."[58]

Weapon focus effect (WFE)

Latest research

Weapons do not *increase false IDs.* A 2016 meta-analysis found that, because there were only four studies on target absent arrays, which had contradictory results, "[n]o conclusions can be drawn regarding the WFE on TA [target absent] lineups . . . . This also implies that there seems not to be sufficient evidence for experts testifying for the defense on the weapon focus effect with respect to identification decisions, as this type of expert testimony typically focuses on factors contributing to the likelihood of false identifications."[59]

---

[54] Egeth, *Emotion and the Eyewitness*, 245, 252 (Academic Press 1994), *citing* Hotch & Bothwell, *Arousal, description and identification accuracy of victims and bystanders*, J. Social Behavior & Personality, 481 (1990)).

[55] Herve, Cooper & Yuille, *Biopsychosocial Perspectives on Memory Variability in Eyewitnesses* in APPLIED ISSUES IN INVESTIGATIVE INTERVIEWING, EYEWITNESS MEMORY AND CREDIBILITY ASSESSMENT, 99, 106 (Springer 2013) ("Laboratory-based methods are, for ethical reasons, unable to evoke remarkable memories as the stimuli used cannot produce extreme stress or trauma.").

[56] Christianson, *Emotional Stress and Eyewitness Memory: A Critical Review*, Psychol. Bull., 284, 297 (1992).

[57] Valentine & Mesout, *Eyewitness Identification Under Stress in the London Dungeon*, Appl. Cog. Psychol., 151 (2008).

[58] *Id.* at 158. *See also* Wixted, Mickes, Dunn, Clark & W. Wells, *Estimating the reliability of eyewitness identifications from police lineups*, Proceedings of the National Academy of Sciences, 304, 309 (2016) (In a field study of mostly cross-race armed robbery cases, which would presumably create high stress in many victims and witnesses, the researchers concluded that "confidence in an eyewitness identification from a fair lineup is a highly reliable indicator of accuracy."). *See* W. Wells, Campbell, Li & Swindle, *The characteristics and results of eyewitness identification procedures conducted during robbery investigations in Houston TX*, Policing, An Internat'l J. Police Strategies & Management, 601, 608, Table II (2016) (74% of the robbers were armed and 62% of the crimes were cross-race.).

[59] Kocab & Sporer, *The Weapon Focus Effect for Person Identifications and Descriptions: A Meta-analysis,* ADVANCES IN PSYCHOLOGY AND LAW, Vol. I, 71, 105 (Miller & Bornstein eds. 2016).

High confidence = high accuracy. One 2016 study found an increase in the false ID rate when a weapon is present but the effect was very small compared to a much larger effect on the overall correct ID rate.[60] For highly confident witnesses, the WFE was negligible. Thus, the authors asked, "can identifications made by highly confident eyewitnesses (those most likely to make it to trial) be trusted? In other words, are these identifications highly accurate? [The data] show that they are."[61]

Older research

No WFE in field and archival studies. Neither field nor archival studies have found a weapon focus effect.[62]

The WFE is small in the laboratory. The weapon focus effect on identification accuracy is small, 10%. This reduction has been described as "small" or "not of great magnitude."[63]

The WFE disappears if exposure is long enough. The WFE is said to be a reliable effect "particularly in crimes of short duration" when the weapon is visible.[64] However, there is no "systematic exploration of exposure duration" in the WFE literature.[65]

There is no WFE when the witness sees the culprit before the weapon. One study found no weapon focus effect on memory for information when the perpetrator was seen before he displayed a gun.[66]

There is no WFE for close proximity. In an effort to explain why various field studies have not found a significant weapon focus effect, researchers have theorized that "close" proximity to the armed perpetrator may reduce the effect.[67]

Cross-race

Latest research

Two studies on the cross-race effect have been published recently. In one, there was a difference of only 3% between same- and cross-race faces (black and white) in the aggregate (all

[60] Carlson, Dias, Weatherford & Carlson, *An Investigation of the Weapon Focus Effect and the Confidence-Accuracy Relationships for Eyewitness Identification*, J. Appl. Res. Mem. & Cog., 1, 1, 6 (2016).
[61] *Id*. at 7.
[62] Behrman & Davey, *Eyewitness Identification in Actual Criminal Cases: an Archival Analysis,* Law & Hum. Behav., 475, 487 (2001).
[63] Steblay, *A Meta–Analytic Review of the Weapon Focus Effect,* Law & Hum. Behav., 413, 417, 420 (1992).
[64] Steblay (1992), *supra,* at 421.
[65] Fawcett, Russell, Peace & Christie, *Of Guns and Geese: A Meta-Analytic Review of the 'Weapon Focus' Literature*, Psychol., Crime & L., 35, 43 (2013); *id.* at 56
[66] *See* Mitchell, Livosky & Mather, *The Weapon Focus Effect Revisited: The Role of Novelty*, Legal & Criminol. Psychol., 291, 295 (1998) ("This finding supports the contention that the obtained [reduction] is in fact an encoding phenomenon occurring as a function of exposure to, and affecting only details seen at the same time as, the experimental item.").
[67] Fawcett, Russell, Peace & Christie, *Of Guns and Geese: A Meta-Analytic Review of the 'Weapon Focus' Literature*, Psychol., Crime & L., 35, 44 (2013) (citations omitted).

confidence levels combined) and a 3% difference for witnesses who were 100% confident.[68] These small differences may be statistically significant, *see* Glossary, but the magnitude of the effects are not substantive enough to conclude that as a general matter people may be less accurate in identifying a person of another race. The other study concluded that a high-confidence initial identification was equally trustworthy for same and cross-race identifications.[69]

Older research

No cross-race effect in field and archival studies. Neither field nor archival studies have found a cross-race effect.[70]

Small differences. Earlier laboratory studies have reported small differences between same- and cross-race identifications. In one study, there was, at most, a three-point difference between same- and cross-race accuracy, and as little as no difference: 79-76, 80-80, 76-73 and 75-74.[71] In another study, the authors wrote, "[a]lthough the cross-race effect is known to be a robust effect, the effect sizes reported here are rather small; thus, replication of these results is necessary."[72] Other studies have reported larger differences (up to about 15%). Without scaling the results by confidence levels, however, it is difficult to tell whether there is any cross-race difference for witnesses who are likely to testify at trial.

Facial recognition studies. The cross-race effect has been studied mostly in facial recognition studies where many faces are seen for very short amounts of time.[73] *See* Glossary. It is not clear, then, that the results apply to real eyewitnesses to crime who usually see only one or two faces (even if there are more perpetrators) for longer amounts of time.

Small/tiny exposure durations. At an exposure duration of as little as 1 second compared to 1/10 of a second, cross-race effects "were not observed."[74]

---

[68] Dodson & Dobolyi, *Confidence and Eyewitness Identifications: The Cross–Race Effect, Decision Time and Accuracy,* Appl. Cog. Pyschol., 113, 118, Figure 1 (2016). The data in this study were reanalyzed by Wixted & G. Wells (2017), *supra,* at 32, Figure 4F, who found virtually no difference in accuracy for same-and cross-race identifications when witnesses are 100% confident and a very small difference in accuracy when witnesses are 80% confident.

[69] Nguyen, Pezdek & Wixted, *Evidence for a confidence-accuracy relationship in memory for same- and cross-race faces*, Q. J. Exp. Psychol., 1 (2016).

[70] Behrman & Davey, *Eyewitness Identification in Actual Criminal Cases: an Archival Analysis,* Law & Hum. Behav., 475, 487 (2001) ("none of the classic eyewitness factors, race, weapon presence or witness type, produced significant or even marginally significant effects when the identifications were made at field showups.").

[71] Marcon, Meissner, Fruch, Susa & MacLin, *Perceptual identification and the cross-race effect*, Visual Cognition, 767, 771 (2010) (Hispanic and African American).

[72] Jackiw, Arburthnot, Pfeifer, Marcon & Meissner, *Examining the Cross-Race Effect in Lineup Identification Using Caucasian and First Nations Samples,* Canadian J. Behavioural Science, 52, 56 (2008) (First Nation and White).

[73] *See, e.g*., Meissner & Brigham, *Thirty Years of Investigating the Own-Race Bias in Memory for Faces; A Meta-Analytic Review*, Psychol. Pub. Pol'y & L., 3, 19 (2001) (91% of the articles were facial recognition studies).

[74] Marcon, Meissner, Fruch, Susa & MacLin, *Perceptual identification and the cross-race effect*, Visual Cognition, 767, 771-772 (2010) ("significant [cross-race effects] were observed at the 100 ms and 500 ms encoding conditions, but were not observed when encoding time was 1000 ms and 1500 ms (statistical formulas omitted)."). *See* MacLin, MacLin & Malpass, *Race Arousal, Attention, Exposure, and Delay: an Examination of Factors*

Cross-race contact. There are mixed conclusions about whether the amount of contact with members of another race alone eliminates the cross-race effect, but it is clear that the cross-race effect does not apply when "it is important to differentiate between individuals in the category in the course of everyday life," such as parents, bosses, and other influential social contacts."[75]

Angry faces. "[C]ross-race faces displaying expressions of anger, a biologically prepotent facial expression that motivates attentional scrutiny and accurate memory[,] eliminate the own-race bias."[76] Moreover, this is true whether or not the participants had significant contact with people of the other race. Because perpetrators are likely to appear angry, the cross-race effect is not likely to apply to violent crimes.

Moving faces: In real life, witnesses see moving faces. In most cross-race studies, faces are static. One study that found that the "significant" cross-race effect present in the single static pose disappeared when the faces were moving.[77]

Distance and cross-race. In a field study primarily examining distance, the authors "found no evidence for significant effects of [cross-race] on identification performance."[78]

Opportunity to view

Distance and lighting (along with duration, delay, disguise, intoxication, and description) are factors with which the people have experience in their own lives. There are few studies in these areas (which are not necessarily consistent with each other), and would add little or nothing to a citizen's own appreciation of how differences may affect one's ability to identify a person.

Distance/Lighting

Older research

*Distance and lighting***.** The Special Master in Henderson was under the impression that "faces are essentially unrecognizable at 134 feet."[79] This is not consistent with the research. One study shows, for example, at a distance of 131 feet and illumination of 3000 lux (cloudy day), 29% correctly identified a stranger and 9% incorrectly identified someone else.[80] "Further research

---

Moderating Face Recognition, 7 Psychol. Pub. Pol'y, 134, 135-136 (2001 (0.5 to 5 seconds); Meissner & Brigham (2001), *supra*, at 19, (median exposure duration of 3 seconds).

[75] MacLin & Malpass, *Racial Categorization of Faces, The Ambiguous Race Face Effect,* Psychol. Pub. Pol'y & L., 98, 99-100 (2001).

[76] Young & Hugenberg, *Individuation Motivation and Face Experience Can Operate Jointly to Produce the Own-Race Bias*, Social Psychological & Personality Science, 80, 83, 84 (2012) (citations omitted). *Id*. at 84 (citations omitted).

[77] Zhao, Hayward & Bulthoff, *Face format at encoding affects the other-race effect in face memory*, J. Vision, 1, 5, 6 & 7 (2014).

[78] Lindsay, Semmler, Weber, Brewer & Lindsay, *How Variations in Distance Affect Eyewitness Reports and Identification Accuracy*, Law & Hum. Behav., 1, 8 (2008) ("[i]f the cross-race effect is robust, one would have expected it to occur under the conditions of our study.").

[79] Report of the Special Master, *State v. Henderson*, at 45.

[80] Wagenaar & Van der Schrier, *Face Recognition as a Function of Distance and Illumination: A Practical Tool for Use in the Courtroom*, Psychol., Crime & L., 321, 328, Table 2 (1996). Lindsay, Semmler, Weber, Brewer & Lindsay, *How Variations in Distance Affect Eyewitness Reports and Identification Accuracy*, Law & Hum. Behav., 1, 8, 9

would be required to determine the distance and conditions that reduce diagnosticity to one and thus probative value to zero."[81]

At the other end of the spectrum, the smallest distance studied is 3 meters, or about 10 feet, so there is no research on identification accuracy for shorter distances.

Exposure time/Retention interval

Latest research

Short exposures decrease the proportion of high-confidence IDs, but not their accuracy. In a study comparing 5-second and 90-second exposures, "accuracy clearly increased with confidence in all exposure and retention interval conditions. This was particularly evident in the upper half of the confidence scale, and especially at the upper end of the scale (i.e., 90%-100% vs. 70%-80%) confidence). Together, these results suggest . . . confidence remained a useful indicator of accuracy in all experimental conditions."[82]

Reanalyzing the same data using *suspect* identifications only, researchers found that "not surprisingly, memory was better (discriminability was higher) when exposure duration was longer . . . ." The data "indicate that participants appreciated the effect that exposure time would have on their memory and compensated for it by appropriately adjusting their confidence, particularly at the high-confidence end of the scale. . . . [Thus,] a high-confidence ID made from the 5 [second] condition was as likely to be correct as a high-confidence ID made from the 90 [second] condition."[83] The key point here is that "while participants in [the 5 second exposure] condition were less likely to make relatively high-confidence IDs, when they did, they were as accurate as high-confidence IDs from the long [90 second] exposure condition."[84]

In the leading meta-analysis of exposure duration, subjects were more accurate for "long" exposures than "short ones," but the median difference between them was only 4.7 seconds and for half of the studies, the "long" exposures ranged from 1.25 seconds to 6 seconds.[85]

(2008) ("[e]ven at 43 m [141 feet], identification evidence has some diagnostic value, and therefore probative value as well.").

[81] Lindsay, Semmler, Weber Brewer & Lindsay, *How Variations in Distance Affect Eyewitness Reports and Identification Accuracy*, Law & Hum. Behav., 1, 8 (2008).

[82] Palmer et al. (2013), *supra*, at 61

[83] Mickes, *Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory*, J. Appl. Res. Mem. & Cog., 93, 96 (2015).

[84] Mickes (2015), *supra,* at 96.

[85] Bornstein, Deffenbacher, Penrod & McGorty, *Effects of exposure time and cognitive operations on facial identification accuracy: a meta-analysis of two variables associated with initial memory strength*, Psychol., Crime & L., 473, 477 (2012). The difference in exposure times for the 16 studies below the median ranged from 0.7 to 4 seconds. The difference in exposure times for the 16 studies above the median ranged from 10 seconds to one hour. *Id.* at 478, Table 1. Extrapolating from the data on Table 1, the ranges below the median would have been along the lines of: 1.25 to 3.75 seconds (Wallace); 2 to 4 seconds (Brigham); 1 to 5 seconds (Malpass); 3 to 6 seconds (Meissner); 0.2 to 1.5 seconds (Semmler & Brewer). The longer exposure durations were 1.5 seconds, 3.75 seconds, 4 seconds, 5 seconds, and 6 seconds. This amount of time might be described as "brief or fleeting," yet identification accuracy was better than at the shorter exposure durations.

Long retention intervals decrease the proportion of high-confidence identifications but not their accuracy. With respect to retention interval, researchers reanalyzed data from four earlier studies to conclude that "high-confidence accuracy remained extremely reliable even as memory conditions deteriorated, so much so that high-confidence suspect ID accuracy was close to 100% correct whether the retention interval was as short as 1 week or as long as 9 months."[86]

Intoxication

Latest Research

At least moderate alcohol consumption has little or no negative effect on eyewitness accuracy, and may improve it. A 2016 article found that "intoxicated participants were no less likely than sober or placebo participants to make an accurate identification from a TP [target-present] lineup," and there was "no significant association between alcohol condition and identification decision in a TA [target-absent] lineup."[87] A 2013 study found that "intoxicated eyewitnesses performed on the same level as their sober counterparts."[88]

Older research

Alcohol does not increase false identifications. A 2007 study found that "[o]n recognition memory tasks, alcohol has been shown to decrease hit rates . . . but to have no effect on false alarm rates."[89]

In a study comparing both "intoxication" and cross-race identification, those in the alcohol condition made fewer identifications – but alcohol diminished the difference between accurate same and cross-race identifications to 2 percentage points (81% v. 79%, "a small tendency,") compared to 5 percentage points in the non-alcohol condition (86% v. 81%, "significantly higher").[90]

Descriptions

[86] Wixted, Read & D. Lindsay, *The Effect of Retention Interval on Eyewitness Identification Confidence-Accuracy Relationship*, J. Appl. Res. Mem. & Cog., 1, 9 (2016).

[87] Kneller & Harvey, *Lineup Identification accuracy: The effects of alcohol, target presence, confidence ratings, and response time*, European J. Psychol. Appl. to Legal Context, 11, 15 (2016). In fact, 45% of the subjects in the alcohol condition rejected the TA lineup compared to 50% in the placebo condition and 40% in the control (no-alcohol) condition.

[88] Hagsand, Roos-af-Hjelmsater, Granhag, Fahlke & Soderpalm-Gordh, *Bottled memories: On how alcohol affects eyewitness recall*, Scandinavian J. Psychol., 188, 193 (2013). The results actually could be interpreted to suggest that alcohol increases eyewitness performance. In target present lineups, 40% in the higher alcohol dose group correctly identified the target compared to 5% in the lower alcohol dose group and 25% in the no alcohol control group. Similarly, in target absent lineups, 45% of the higher alcohol dose group correctly rejected the lineup compared to 36.4% in the lower alcohol dose group and 23.8% in the no alcohol control group. *Id.* at 37, Table 1 and 38, Table 2.

[89] Mintzer, *The acute effects of alcohol on memory: A review of laboratory studies in healthy adults*, Int. J. Disabil. Hum. Dev., 397, 399 (2007) (parentheticals omitted).

[90] Hilliar & Kemp, *Now Everyone Looks the Same: Alcohol Intoxication Reduces the Own-Race Bias in Face Recognition*, Law & Hum. Behav., 367, 372, 373, Figure 2 (2010). The entire span of accurate identifications regardless of alcohol consumption or racial differences in this study was seven percentage points (79% to 86%).[90]

New research

Witnesses are better at identifying targets than describing them. "Despite the clear intuition that witnesses who are better at describing a target should also be better at recognizing him, this relationship has proved to be quite elusive and generally weak."[91] Moreover, "efforts to describe a previously seen face can actually *impair* subsequent memory performance, at least under some circumstances."[92]

Disguise

Research on the effect of disguises is not consistent. In some studies, a hat reduced identification accuracy; in others, it did not. In at least two experiments, a hat did not increase false identifications although it reduced true ones.[93] In another, "[t]arget disguise [baseball cap and dark sunglasses] . . . had no significant main effects on identification in either the target-present or the target-absent lineups."[94]

Overall, confidence appears to decrease with disguises, indicating that "witnesses are sensitive to some degree of their ability to make accurate identifications, and this sensitivity is in a form (confidence) understood by triers of fact."[95]

[91] Meissner, Sporer & Schooler, *Person Descriptions as Eyewitness Evidence*, in Lindsay, Ross, Read & Toglia, eds. THE HANDBOOK OF EYEWITNESS PSYCHOLOGY, Vol. II, 3, 21 (2014).
[92] Meissner et al. (2014), *supra*, at 22; *id*. at 21 ("the elicitation of elaborate verbal descriptions may lead participants to generate inaccurate details which then impairs their recognition performance").
[93] Mansour, Beaudry, Bertrand, Kalmet, Melsom & Lindsay, *Impact of Disguise on Identification Decisions and Confidence with Simultaneous and Sequential Lineups*, Law & Hum. Behav., 514, 518 Table 1 (2012) *id.* 521. Table 3.
[94] Yarmey, *Eyewitness Recall and Photo ID: A Field Experiment*, Psychol., Crime & L., 53, 65 (2004).
[95] Mansour et al. (2012), *supra,* at 524.

SUMMARY
CONFIDENCE AND ACCURACY

> *The idea that initial confidence is not strongly related to accuracy conflicts with virtually all empirical evidence – both in the lab and in the real world – that has accumulated over the last 15 to 20 years. . . . These studies have established beyond any reasonable doubt that, for adults who make an ID from a lineup, the relationship between initial confidence and accuracy in a typical forensically-relevant lab study – precisely the kind of study that once convinced the field that the relationship is weak – is in fact strong."*

Wixted, Mickes, Clark, Gronlund & Roediger III, *Initial Eyewitness Confidence Reliably Predicts Identification Accuracy*, American Psychological Assn., 515, 516, 518 (September 2015) (emphasis in original).

In the criminal justice system, the real issue is the misidentification of an innocent suspect as the perpetrator. Fillers (foils) are known to be innocent and are not prosecuted; guilty suspects who are not identified are not prosecuted – at least not on the basis of a non-identification. Standing alone, suspects who are identified with less than high confidence, are not prosecuted. Thus, the focus should be on the accuracy of witnesses who identify *suspects* at different levels of confidence.[96]

Statistical methods that

- include fillers and non-identifications, and
- do not plot levels of confidence for suspect identifications

do not provide factfinders with the information they need to assess how accurate a witness is who has identified the suspect with a given level of confidence.

The most recent research on the relationship between confidence and accuracy consistently shows that there is high correspondence between confidence and accuracy for suspect identifications. Analyzing new data and reanalyzing data from earlier studies to focus on suspect identifications only, researchers have found that subjects who say they are 90-100% confident are 95-100% accurate. One study found that "even low-confidence [0-60%] suspect IDs are fairly likely to be correct (about 83% correct), though most would probably agree that the 17% error rate is too high to justify a conviction based on a low-confidence ID alone."[97]
Earlier research that reached a different conclusion based on different data and/or different statistical methods should now be disregarded.

[96] This does not mean that filler or non-IDs are unimportant. At an earlier stage in the criminal justice process, they may direct the police away from a particular suspect. But at the time of trial, the critical issue is how reliable a witness is who picked the suspect. Filler and non-identifications do not bear on that issue.
[97] Wixted et al. (2015), *supra*, at 519. *See* Wixted & G. Wells, *The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, Psychological Science in the Public Interest, 10, 30 and 37, Figure 5A (2017). Accuracy rates for subjects and witnesses in the lowest confidence category (0-20%) varies considerably around the 64% mark with some outliers considerably above and below that percentage, but accuracy does not vary much for witnesses in the highest categories (94.2-99.7% accuracy for those who are 90-100% confident, and 90.3% accuracy for those who are 70-80% confident).

CONFIDENCE AND ACCURACY

> *The idea that initial confidence is not strongly related to accuracy conflicts with virtually all empirical evidence – both in the lab and in the real world – that has accumulated over the last 15 to 20 years. . . . These studies have established beyond any reasonable doubt that, for adults who make an ID from a lineup, the relationship between initial confidence and accuracy in a typical forensically-relevant lab study – precisely the kind of study that once convinced the field that the relationship is weak – is in fact strong.*

Wixted, Mickes, Clark, Gronlund & Roediger III, *Initial Eyewitness Confidence Reliably Predicts Identification Accuracy*, American Psychological Assn., 515, 516, 518 (September 2015) (emphasis in original).

Research as Summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized the research on the relationship between a witness's confidence and his or her accuracy as follows: "A witness's level of confidence, standing alone, may not be an indication of the reliability of the identification. Although some research has found that highly confident witnesses are more likely to make accurate identifications, eyewitness confidence is generally an unreliable indicator of accuracy."[98]

Analysis

Recent studies indicate that the opposite of what *Henderson* says is true, that is: an eyewitness's level of confidence at the time of the initial identification is a highly reliable indicator of accuracy, with accuracy rates for high-confidence witnesses typically ranging from 95-100% and for moderate- and low-confidence witnesses 90% and 73% respectively. The most recent article on this subject finds that given an initial uncontaminated memory test using fair lineups, with no lineup administrator influence, and an immediate confidence statement, "mock-crime studies and police department field studies consistently show that, for adults, (1) confidence and accuracy are strongly related and (2) high-confidence suspect IDs are remarkably accurate."[99] This adds to the "growing number of studies [showing that] the magnitude of an individual's confidence rating in a lineup decision can be well calibrated with its likely

[98] Defense experts have led courts to believe that the relationship between confidence and accuracy is weak. *See, e.g., Thomas v. United States*, 59 A.3d 1252, 1266 (D.C. 2013) ("In *Benn* [*v. United States*, 978 A.2d 1257 (D.C. 2009)], we recognized that '[r]esearch reveals ... that the correlation between a witness's expression of certainty in an identification and its accuracy is, at a minimum, greatly overstated, and perhaps unwarranted,' *id.* at 1268, and the 'correlation between confidence and accuracy of an identification ... are counterintuitive.' *Id.* at 1277; *see also Hager v. United States,* 856 A.2d 1143, 1148, amended on other grounds, 861 A.2d 601 (D.C. 2004) ('the correlation between witness confidence and accuracy ... may well be beyond the ken of the average layperson.').").

[99] Wixted & G. Wells, *The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, Psychological Science in the Public Interest, 10, Abstract (2017).

accuracy,"[100] An older mixed archival and laboratory study also concluded that "witnesses who display high levels of certainty . . . are unlikely to choose innocent persons."[101]

Moreover, there appears to be consensus that, while there is no clear demarcation, the speed of a decision is indicative of accuracy. "Fast identifications (6 seconds or less) and confident (90-100%) individuals showed an impressive 97% accuracy rate when they selected someone from a lineup."[102] Another set of researchers observed, "a striking feature of our data is that impressive accuracy rates were obtained when high confidence and the 10-s boundary were used together as a marker of accuracy for adult samples. Specifically, the combined use of 90-100% confidence and the 10-s time boundary diagnosed identification decisions with a high probability of accuracy (88.1%) overall."[103] Jump out identifications (without regard to levels of confidence) have been described as "desirable witness expressions of absolute certainty."[104]

These studies alone make it inappropriate to claim that "eyewitness confidence is generally an unreliable indicator of accuracy." At best, this claim is disputed; at worst, this claim is flat out wrong. Making recommendations about the "unreliability of eyewitness identification" under such circumstances is unwarranted.

The Confidence-Accuracy Characteristic establishes that highly confident suspect identifications are highly accurate and moderately confident witnesses are only slightly less so.

---

[100] Dobolyi & Dodson, *Eyewitness Confidence in Simultaneous and Sequential Lineups,* J. Exp. Psychol.: Appl., 1 (2013) (citations and parentheticals omitted). Well-calibrated means that the level of confidence accords with the level of accuracy.

[101] Behrman & Richards, *Suspect/Foil Identification in Actual Crimes and in the Laboratory: A Reality Monitoring Analysis*, Law & Hum. Behav., 279, 297 (2005). *See also* Palmer, Brewer, Weber & Nagesh, *The Confidence-Accuracy Relationship for Eyewitness Identification Decisions: Effects of Exposure Duration, Retention Interval and Divided Attention*, Am. Psych. Assn., 55, 56 (2013) ("[T]here is a growing body of evidence that points to a meaningful CA relationship under certain conditions. For example, confidence has been shown to be a useful predictor of accuracy when . . . the witness makes a positive identification as opposed to rejecting the lineup."); Lindsay, Read & Sharma, *Accuracy and Confidence in Person Identification, the Relationship is Strong When Witnessing Conditions Vary Widely,* Am. Psychol. Soc., 215, 217 (May 1998) ("These considerations [variability in exposure duration, retention interval, unusual appearance, difference in appearance] suggest that ability to identify the perpetrator varies greatly from one real-world witness to another, and thus, in turn, suggests that the real-world AC relationship is strong.").

[102] Dodson & Dobolyi, *Confidence and eyewitness identifications: The cross-race effect, decision time and accuracy,* Appl. Cog. Psychol., 113, (2016), *citing* Sauerland & Sporer, *Fast and confident: postdicting eyewitness identification in a field study*, J. Exp. Psychol.: Appl., 646 (2009).

[103] Weber, Brewer, Wells, Semmler & Keast, *Eyewitness Identification Accuracy and Response Latency: The Unruly 10-12-Second Rule*, Am. Psychol. Assn., 139, 146 (2004). *Accord* Dunning & Perretta, *Automaticity and eyewitness accuracy: a 10-to-12 second rule for distinguishing accurate from inaccurate positive identifications*, J. Appl. Psychol., 951, 951 (2002) (Identifications made within 10-12 seconds were 90% accurate compared to 50% for those that took more time.). These studies did not use the Confidence Accuracy Characteristic discussed below. If they had, the accuracy rate would undoubtedly be higher. *See also* Valentine, Pickering & Darling, *Characteristics of Eyewitness Identification that Predict the Outcome of Real Lineups,* Appl. Cog. Psychol., 969, 984 (2003) ("[f]ast decisions are more likely to result in identification of the suspect (87%) than average or slow decisions (38% and 31% respectively)."

[104] Klobuchar, Steblay & Caligiuri, *Improving Eyewitness Identifications: Hennepin County's Blind Sequential Lineup Pilot Project*, Cardozo Public Law, Policy & Ethics J., 381, 399-400 (April 2006).

The real concern in the criminal justice system is the misidentification of an innocent suspect as the perpetrator. In laboratory studies, researchers use "target present" (TP) and "target absent" (TA) arrays. When calculating accuracy, the researchers have often examined correct identifications, correct rejections (in TA arrays), incorrect rejections (in TP arrays), and misidentifications of either a person in a TA array, or a foil or filler in a TP array. In real life, however, (1) true perpetrators are not at risk if they are not identified when they are present in the array (incorrect rejections); (2) innocent suspects are not at risk if they are not identified in an array in which the true perpetrator is absent (correct rejections); and (3) foils or fillers are not at risk because they are known errors (misidentification). All three of these categories, therefore, should be excluded from accuracy rate calculations. This leaves for inclusion in the calculation only correct identifications of the guilty suspect in a TP array and misidentifications of the innocent suspect in a TA array. This is not a novel concept. In 1980, "Wells and Lindsay . . . defined the diagnosticity ratio as the ratio of (the rates of) correct to false suspect identifications, [asserting that t]he higher the ratio, the more likely an identified suspect is to be the criminal." [105]

The concept of suspect identifications was developed further in 2015 when a researcher coined the phrase "confidence-accuracy characteristic" [CAC] that "simply consists of plotting identification accuracy of suspect IDs (ignoring filler IDs) for each level of confidence regardless of the specific scale that is used."[106] What this comes down to is dividing the number of correct *suspect* identifications at each level of confidence by the number of correct plus false *suspect* identifications (often estimated by dividing all filler identifications from fair target-absent lineups by the number in the array) to arrive at the proportion of correct identifications.

In some studies, researchers designate one member of a target-absent lineup as the suspect, often choosing a person who closely resembles the target, and the rest are fillers.[107] In such studies, researchers would use only the number of subjects who chose the designated innocent suspect to compute the false identification rate. In the majority of studies, however, no member of the target-absent lineup is designated as the innocent suspect (i.e., all of the lineup members are fillers). As is well known in the field of eyewitness identification, in order to properly estimate incorrect suspect IDs from fair target-absent lineups in the latter condition, one should divide filler identifications in a target-absent array by "n", where "n" is the size of the array.[108] "Although this approach involves treating target-present and target-absent lineups differently, it is – of all logistically feasible methods – the most stringent test of confidence as an indicator of identification accuracy."[109]

---

[105] R. Lindsay, Semmler, Weber, Brewer & M. Lindsay, *How Variations in Distance Affect Eyewitness Reports and Identification Accuracy*, Law & Hum. Behav., 1, 8, 9 (2008), *citing* Wells & Lindsay,*n estimating the diagnosticity of eyewitness nonidentifications*, Psychol. Bull., 776 (1980).

[106] Mickes, *Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory*, J. Appl. Res. Mem. & Cog., 93, 101 (2015).

[107] In such studies, only identifications of the designated innocent suspect would be counted as a misidentification.

[108] *See, e.g.,* Sauer, Brewer & Wells, *Is there a magical time boundary for diagnosing eyewitness identification accuracy in sequential line-ups?*, Legal & Criminol. Psychol., 123, 130 (2008) (Sauer, et al. (2008) ("we assumed that all line-up members are equally like to be the innocent suspect and divided the false identification rates in the target-absent line-ups by 8."); Palmer et al. (2013), *supra*, at 62 ("we estimated the innocent suspect false identification rate for target-absent lineups by dividing the foil identification rate by the number of lineup members (i.e., eight).").

[109] Palmer et al. (2013), at 62. *See also* Clark, Moreland & Gronlund, *Evolution of the empirical and theoretical foundations of eyewitness identification reform*, Psychon. Bull. Rev., 251, 253 (2013) (Some experiments "report only the total identification rate for innocent-suspect lineups, without distinguishing between a filler identification

"[A] CAC plot provides the information that judges and juries want to know when they are trying to assess the reliability of an eyewitness who identified a suspect from a lineup."[110] It answers the question "given that an eyewitness has a particular level of confidence in his/her ID of a suspect, how accurate is that ID likely to be?"[111] Two studies suggest that the answer differs from what has long been assumed:

- Wixted et al. (2015), *supra*, at 519-520: reanalyzing the Palmer data by excluding the known errors of foil identifications, high-confidence witnesses were 98% accurate, moderate-confidence witnesses were 94% accurate, and low- confidence witnesses were 83% accurate;[112] and

- Sauer, Brewer, Zweck & Weber, *The effect of retention interval on the confidence-accuracy relationship for eyewitness identification*, Law & Hum. Behav., 337, 342 (Table 1) (2010): excluding foil identifications and dividing the number of false identifications by the number in the array, highly confident subjects were 97+% accurate in both the immediate and delayed conditions.[113] The accuracy rates in the delay (20-50 days) condition were 47% for 0-20% confidence; 78% for 30-40% confidence; 84% for 50-60% confidence; and 93% for 70-80% confidence, all remarkably high.[114]

In a variety of studies recently reanalyzed using the CAC, accuracy rates for highly confident witnesses were 100% (Read), 95-100% (20 studies), and 96.1% to 98.4% (15 studies).[115] Even those who were not highly confident achieved impressive accuracy rates: 90% for those who were 70-80% confident, 81% for those who were 50-60% confident, and even 64% for those who were 0-20% confident (9 studies).[116]

Thus, correctly analyzing the data yields a very different picture about the relationship between confidence and accuracy and almost uniformly leads to the conclusion that the relationship between confidence at the time of the initial identification and accuracy is strong and that the witnesses who are most likely to testify at trial are under-confident rather than over-confident. Moreover, confidence appears to outweigh various system and estimator variables that are said to adversely affect eyewitness

---

and the false identification of an innocent suspect. For those experiments, the false identification rate can be estimated by dividing the total identification rate by the number of people in the lineup."); *see also* Wixted & G. Wells (2017), *supra*, at 20.

[110] Wixted & G. Wells (2017), *supra*, at 24.

[111] Wixted & G. Wells (2017), *supra*, at 56.

[112] "[E]ven low-confidence [0-60%] suspect IDs are fairly likely to be correct (about 83% correct), though most would probably agree that the 17% error rate is too high to justify a conviction based on a low-confidence ID alone." *Wixted et al.* (2015)*, supra,* at 519. *See In re As.H*., 851 A.2d 456 (2004) (20-30% uncertainty, standing alone, is reasonable doubt as a matter of law) Of course, if there is other evidence of identity, even a weak eyewitness identification could be considered by the jury in assessing whether guilt has been proven beyond a reasonable doubt.

[113] Although not discussed in text, the diagnosticity ratios in Table 1 for highly confident choosers of 37.79 (immediate) and 20.47 (delayed) translate to 97% and 95% accuracy.

[114] These percentages are not in the article but can be computed by applying the CAC to the data in Table I for immediate and delayed choosers who made correct identifications and false identifications. In the immediate condition, the accuracy rates were considerably higher except for the two highest confidence level: 81.6% for 0-20% confidence; 88% for 30-40% confidence; 91% for 50-60% confidence; 95% for 70-80% confidence (compared to 93% for delay); and 97.5 for 90-100% confidence (compared to 97% for delay).

[115] Wixted & G. Wells, *supra*, at 27, Figure 3, 30, Figure 4A-S.

[116] Wixted & G. Wells, *supra*, at 37, Figure 5A.

identification. It turns out that these factors may reduce the proportion of high- or moderate-confidence identifications but they do not appreciably diminish the accuracy of those that remain.

In real life, researchers have found an "impressive relationship" between confidence and accuracy in archival and field studies.

Historical reasons for misunderstanding the strength of the relationship between confidence and accuracy.

An understanding of how we got to this point may be useful. There are essentially five reasons that contribute to the misimpression that eyewitnesses are not reliable:

(a) Some early research combined choosers and non-choosers in analyzing the data, thereby skewing the results. Only choosers who identify suspects end up testifying against those suspects at trial.

(b) Researchers have used, and continue to use, a statistical method called point-biserial correlation that is misleading on the issue juries are asked to assess: how closely a stated level of confidence accords with accuracy, especially for witnesses who testify at trial. The point-biserial correlation is a misleading statistic in this context.[117] Using a more straightforward approach, there is a strong relationship between confidence and accuracy.

(c) Researchers counted foils/fillers in their calculations, even though witnesses who choose a filler do not testify at trial; eliminating foil/filler picks and focusing solely on suspect picks (true or false) – the only picks that matter in a trial – reveals a strong relationship between confidence and accuracy.

d) The "consensus" that an eyewitness's confidence is not a good predictor of his or her identification accuracy was based a single survey of a small number of experts who may not have been representative of the field and who, themselves, may have relied on research using the point-biserial correlation coefficient.

e) There is confusion about the reliability of the level of confidence expressed by the witness at the time of the confrontation (initial identification procedure) and the reliability of a different and higher level of confidence at trial.

(a) Choosers and non-choosers

Researchers initially included both choosers and non-choosers in their calculations, only to then determine that including non-choosers skewed the results. "It appears that the counterintuitive finding – confidence is not a good predictor of identification accuracy – stressed by many researchers and psychological experts in their courtroom testimony may only characterize broad comparisons of witnesses (i.e. including witnesses who make positive identifications as well as witnesses who reject lineups). These present findings indicate that when limited to witnesses who make positive

[117] *See* Juslin, Olsson & Winman, *Calibration and Diagnosticity of Confidence in EWID: Comment on What Can Be Inferred From the Low Confidence-Accuracy Correlation,* J. Exp. Psychol., 1304, 1305 (1996); Wixted & G. Wells (2017), *supra*, at 49; and discussion below.

identifications under laboratory conditions [versus witnesses who do not make any identification at all], confidence appears to be a somewhat stronger predictor of accuracy."[118] The relationship between confidence and accuracy for people who do not pick anyone in an array (non-choosers) is poor, the reason for which "remains unclear."[119] Because non-choosers are not likely to be eyewitness identification witnesses at trial, they will not contribute to a conviction, whether the suspect is actually guilty or actually innocent. Their accuracy rates, therefore, are not relevant to a jury determination. Including non-choosers in the calculation results in an inaccurate assessment of the relationship between confidence and accuracy of testifying eyewitnesses.

Indeed, the correlation between confidence and accuracy rose from 0.08 reported by Wells & Murray[120] for choosers and non-choosers combined to .41 for choosers only in this meta-analysis.[121] An effect size of .41 is a little more than halfway between a moderate (.30) and strong (.50) effect size.[122] So using this statistical method, called point-biserial correlation coefficient, it would be inappropriate to characterize the correlation for choosers only as weak to moderate.[123] Indeed, for choosers only, "Lindsay et al. (1998) obtained confidence-accuracy correlations of .51 and .68 for two different videos when the witnessing conditions varied widely."[124] These would be regarded as strong correlations. This has led to the conclusion that "the point-biserial correlation is now known to be quite a bit higher than it was once thought to be."[125]

(b) Point-biserial correlation coefficient[126]

[118] Sporer, Penrod, Read & Cutler, *Choosing, Confidence, and Accuracy: A Meta-analysis of the Confidence-Accuracy Relation in EWID Studies*, Psychol. Bull., 315, 324 (1995); *see* Dobolyi & Dodson, *supra*, at 1 ("although early research indicated that a witness's confidence is usually weakly correlated with his or her accuracy, many of these studies did not distinguish between confidence for 'choosers' . . . and 'non-choosers'"); Sauer et al. (2010), *supra,* at 337 ("Studies using the calibration approach . . . have . . . demonstrated robust CA relationships when participants positively identify a lineup member as the culprit . . . .").

[119] Sauer et al. (2010), *supra*, at 338; *see* Dobolyi & Dodson, *supra*, at 9, Figure 2(b).

[120] Wells & Murray, *Eyewitness Confidence*, in EYEWITNESS TESTIMONY: PSYCHOLOGICAL PERSPECTIVES (Cambridge University Press 1984).

[121] Sporer et al. (1995), *supra*, at 315, 319.

[122] *See* Cohen, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES, 286-288 (2d ed. 1988) (effect sizes: .10 = small; .30 = medium; .50 = large).

[123] Nevertheless, Dr. Penrod has claimed that there is "little if any relationship between the expressed confidence of an eyewitness and the actual accuracy of that eyewitness's identification." Proffer in *United States v. Jacques Parker*, 2012 CF3 1232 (D.C. 4/18/2012), and that "[t]he correlation between confidence and accuracy is not nearly as strong as the average layperson tends to believe." Proffer in *United States v. Anthony Wright*, 2014 CF3 16589 (D.C. 6/16/2015). Professor Brian Cutler has characterized "a witness's confidence in the accuracy of her identification [as] a modest gauge of its accuracy." Proffer in *United States v. Drummond*, 2013 CF3 18138 (D.C. 10/06/14).

[124] Behrman & Davey, E*yewitness Identification in Actual Criminal Cases: An Archival Analysis,* Law & Hum. Behav., 475, 486 (2001).

[125] Wixted et al. (2015), *supra,* at 517, *citing* Lindsay et al. (2008), *supra*.

[126] The point-biserial correlation coefficient is sometimes referred to as "Pearson's r." When one of the two variables is dichotomous (like male vs. female, or correct vs. incorrect) and the other variable is continuous (like height or confidence), the term is "point-biserial correlation coefficient"; when both variables under consideration are continuous (like height and age), the term used is "Pearson's r." But the formulas used to compute both the "point-biserial correlation coefficient" and the "Pearson's r" are exactly the same.

"The notion that scientific research has established the unreliability of eyewitness confidence was largely set in stone in the 1980s and early 1990s when researchers routinely measured the relationship between confidence and accuracy using a potentially misleading statistic – one that is capable of masking (and, as it turns out, actually did mask) the strong relationship that we now know to exist," that is, the point-biserial correlation coefficient. [127] But the point-biserial correlation coefficient is inapposite to what the jury needs to know. What the jury needs to know is: How accurate are witnesses *who testify at trial*? As a general rule, witnesses who do not identify the suspect or who initially identify the suspect with a low level of confidence are not eyewitness identification witnesses at trial, although they may provide other testimony that has a bearing on identification.

The point-biserial correlation does not answer the question of how accurate people are at different levels of confidence at the time of the initial confrontation. Instead,

> [t]he point-biserial correlation compares confidence with the rather unrealistic norm of perfect discrimination; that is[,] all correct identifications should be in one confidence category and all wrong identifications should be in another and lower confidence category. . . . [P]erfect discrimination is not within human capabilities and the factors that constrain memory performance are beyond control of the witness at the time of the confidence assessment.

Juslin et al. (1996), *supra*, at 1305.

"The point-biserial correlation . . . is not very useful information for inferring likely accuracy from confidence."[128] In essence, this statistical method tries to boil down the relationship between confidence and accuracy to a single number. The method aggregates all the correct identifications at all levels of confidence and all the incorrect identifications at all levels of confidence and statistically compares them. A correlation of .41, for example, means that correct identifications are associated with a higher level of accuracy than incorrect identifications. In testimony, Dr. Penrod, one of the experts in *Henderson*, has described a correlation of .41 as meaning that confident witnesses are 50% to 70% accurate. But the statistic itself does not distinguish between witnesses who are 10% or 20% confident – and, therefore, less likely to be accurate – and those witnesses who are 90% or 100% confident – and, therefore, more likely to be accurate. A correlation of .41 for witnesses who chose a photograph in an array does *not* tell us how accurate an individual witness is likely to be. In fact, the "point-biserial correlation may underestimate, or even hide, a useful relation between subjective and objective probabilities of correct identification."[129] Although a high point-biserial correlation coefficient . . . indicates a strong relationship between confidence and accuracy, "a low point-biserial correlation coefficient . . . does not necessarily indicate a weak relationship."[130]

Best practices require that the police obtain a statement of confidence at the time of the initial identification. It does not make sense to do this and then muddy the waters by giving an overall accuracy rate that ignores accuracy at different levels of confidence.

(c) Calibration

[127] Wixted et al. (2015), *supra* at 516.
[128] Palmer et al. (2013), *supra*, at 56.
[129] Juslin et al. (1996), *supra*, at 1314.
[130] Wixted et al. (2015), *supra,* at 517.

Researchers then moved to plotting accuracy at different levels of confidence using a method called "calibration."

> The forensic utility of the calibration approach, when compared to correlation, lies in its indication of a probable accuracy for each level of confidence. As Juslin et al. (1996) note, the knowledge that the CA [confidence-accuracy] correlation is, for example, .28 does not help assess the accuracy of an individual identification made with 80% confidence. On the other hand, knowing that 80% (or 70, or 90%) of identifications made with 80% confidence are correct provides a guide for assessing the likely reliability of an individual identification decision.

Sauer et al. (2010), *supra*, at 338.

In essence, in the context of eyewitness identification, calibration measures "the agreement between subjective assessments of recognition reliability and corresponding objective reliability." [131] As it turns out, confidence at the time of the initial identification pretty closely tracks accuracy, that is, a person who is 100% confident is highly accurate (as described below) and a person who is 30% confident is much less so. If anything, accuracy may exceed confidence along the continuum, that is, a person who is 60% confident may be accurate 70% or more of the time.[132] This is real information that juries can understand and use – and it comes straight from the mouth of the witness without the need for either expert testimony or jury instructions.

There is some variability in the laboratory studies that provide text or charts or tables from which accuracy at various levels of confidence can be ascertained.

- Juslin et al. (1996), *supra*, at 1312: an identification made with 100% certainty implies a probability larger than 94% that the suspect is the culprit;
- Weber et al. (2004), *supra*, at 22: people who are 90-100% confident and pick within 10 seconds are 88.1% accurate;
- Lindsay et al. (2008), *supra*, at 217 (Figure 1): appearing to show only one incorrect identification and 32 correct identifications when the subjects were 100% confident. This means that 97% of people who were 100% confident were also accurate. There were fewer IDs as conditions worsened, but no less accuracy at the 100% mark;
- Palmer et al. (2013), *supra*, at 66 (Figure 3): showing approximately 90% accuracy for persons who said they were 90-100% confident.

[131] Juslin et al. (1996), *supra*, at 1306. Juslin also "showed that the magnitude of the point-biserial correlation can be very low even when the relationship between confidence and accuracy exhibits *perfect* calibration. Perfect calibration exists when the level of confidence expressed by an eyewitness corresponds exactly to the percentage of eyewitnesses who are correct when they express that level of confidence . . . [*e.g.*,] witnesses who express 60% confidence in an ID are correct 60% of the time, and witnesses who express 80% confidence in an ID are correct 80% of the time. Wixted et al. (2015), *supra*, at 517.

[132] We are using percentages here because laboratory studies use numerical scales. In real life, witnesses may volunteer a numerical response, but they are not solicited for one. In fact, they are asked to state how confident they are "in their own words."

Because laboratory studies that use both target-present and target-absent arrays usually divide the subjects equally between them, and in real life, the proportion of arrays with true perpetrators is likely to be higher than those with innocent suspects, the proportion of accurate witnesses is also likely to be higher.[133] Given these studies, it would be false and misleading to suggest that confidence is an unreliable indicator of accuracy.[134]

Although the move toward calibration was a definite improvement over the point-biserial correlation coefficient, it still is not the most appropriate statistical method for assessing accuracy because it includes foils/fillers in its calculations. In other words, because filler identifications are included in the analysis, a calibration plot does not answer the key question concerning the reliability of a *suspect* identification made by an eyewitness. The key question is this: How reliable is a *suspect* identification that has been made by an eyewitness with a particular level of confidence? As described earlier, the answer to that key question is provided by CAC analysis. When CAC analysis is used, the confidence-accuracy relationship turns out to be even more impressive than calibration studies have suggested.

(d) Reliance on outdated and possibly biased data

In 2001, a survey was conducted of eyewitness experts.[135] Given that studies using the point-biserial correlation coefficient dominated the literature at the time, it is not surprising that 87% of the experts thought that the proposition that "an eyewitness's confidence is not a good predictor of his or her

---

[133] *See* Brewer & Wells, *The Confidence-Accuracy Relationship in Eyewitness Identification: Effects of Lineup Instructions, Foil Similarity, and Target Absent Base Rates,* J. Exp. Psychol., 11, 25 (2006) (Figure 3 shows graphically how the relationship between confidence and accuracy rises as the number of target absent arrays diminish). Recently, Gary Wells, who once described confidence as "forensically useless," acknowledged,"calibration tends to be good for those making an identification." Reanalyzing some of his own data, he found "nearly perfect calibration for very high confidence witnesses" when the base rate was 70%. Wells, Yang & Smalarz, *Eyewitness Identification: Bayesian Information Gain, Base-Rate Effect-Equivalency Curves, and Reasonable Suspicion,* Law & Hum. Behav., 99, 118-119 (2015). An archival study comparing DNA results and eyewitness identification found a base rate of at least 68% comparing arrays where DNA proved the perpetrator was present to all the arrays, including those with inconclusive DNA results. Kellstrand, *Eyewitness identification accuracy in cases accepted and rejected for prosecution: An archival analysis of criminal case files*, Unpublished Manuscript, San Diego, University of California (2006). Where DNA established definitively that an array was either perpetrator-present or perpetrator-absent, the perpetrator-present base rate was 95%. *Id.* (cited in Clark & Wells, *On the Diagnosticity of Multiple-Witness Identification*s, Law & Hum. Behav., 406, 416 (2008)). At least in jurisdictions where the police must have an articulable suspicion to place a suspect in an array or show-up (*e.g.*, possession of stolen property, cell site data, license plate numbers, fingerprints, use of a stolen credit card, MO, presence in the area wearing clothing similar to that described by the witness, etc.), the base rate likely would be higher than 50%.

[134] By the same token, in real life, witnesses identify suspects with varying degrees of confidence. It appears that people "can appreciate when conditions are poor and adjust their confidence accordingly," Mickes, *supra*, at 101; *see id.* at 98. A lower level of confidence may be attributable to the fact that the witness did not get a good look at the perpetrator or the suspect is not the perpetrator. Absent other evidence of identification, a prosecution is unlikely to go forward.

[135] Kassin, Tubb, Hosch & Memon, *On the "general acceptance" of eyewitness testimony research: A new survey of the experts*, Am. Psychol., 405 (2001). Sixty-four of the 197 experts to whom a questionnaire was sent returned data in a usable form. *Id*. at 407. It is not clear that this is a representative sample or that the research on which they relied to form their opinions was applicable to real cases.

accuracy" was reliable enough to testify to. *Id.* at 413 (Table 5).[136] Were the same experts presented with new studies discussed above, would their answers be the same?

Moreover, some believe that the survey was not representative of the scientific communities that were considering eyewitness identification issues at the time. One article claimed that "All [the Kassin] studies prove is that those who believe in laboratory-based research as a valid context to study eyewitness memory share some common beliefs."[137] And the study itself acknowledged that "individuals with the most expertise in an area may also have the greatest motivation to present it in a favorable light. . . . This possible confounding of expertise and motivation implies that perhaps our respondents should have been drawn from a broader population of basic experimental psychologists who study non-eyewitness processes or who do not testify in court."[138]

(e) Confidence at the time of the initial identification; misimpression of the exoneration cases

When considering the relationship between confidence and accuracy throughout this discussion, consistent with the Supreme Court's holding in *Manson v. Brathwaite*,[139] we are referring to "the level of certainty demonstrated at the confrontation" (*e.g.*, the time of the initial identification).[140]
The National Academy of Sciences noted that "[i]n many of [the Innocence Project] cases, eyewitness identification played a significant evidentiary role, and almost without exception, the eyewitnesses who testified expressed complete confidence that they had chosen the perpetrator. Many eyewitnesses testified [at trial] with high confidence *despite earlier expressions of uncertainty*."[141] This effect was recognized in *Benn v. United States, supra*, 978 A.2d at 1265), where "looks like" at the time of the initial identification became "absolutely positive," "very sure," "looks just the same" at trial. Thus, although there is support for the proposition that confidence statements given at the time of trial (that are higher than those given earlier) may not be accurate, initial confidence statements remain significant evidence of accuracy.

---

[136] This percentage includes those who "tend to favor" the proposition. Only 45% of those experts who "tend to favor" a proposition would testify. *Id*. (Table 6).
[137] Yuille & Cooper, *Challenging the Eyewitness Expert*, in Ziskin & Faust (Eds.), COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY, 685, 690, (6th edition 2012).
[138] Kassin et al. (2001), *supra*, at 414.
[139] *See Manson v. Brathwaite*, 432 U.S. 98, 114 (1977) (Factors to be considered in assessing reliability "include the opportunity of the witness to view the criminal at the time of the crime, the witness' degree of attention, the accuracy of his prior description of the criminal*, the level of certainty demonstrated at the confrontation*, and the time between the crime and the confrontation.").
[140] We are aware of no field research on how best to take a confidence statement. It is certainly an issue to be explored. Certainty may be evident from the witness's statement alone, taking into consideration, for example, the witness' speed, demeanor, and/or inflection used in making the identification (*e.g.*, "That's him! That's the guy who robbed me."). *See* Douglass, Smith & Fraser-Thill, *A Problem with Double-Blind Photospread Procedures: Photospread Administrators Use One Eyewitness's Confidence to Influence the Identification of another Eyewitness,* Law & Hum. Behav., 543, 549 (2005) (classifying an immediate "That's him! That's the guy!" as a high-confidence identification; and "I think maybe it was him." as a low confidence identification); Behrman & Richards*, supra*, at 284 (classifying "That's him, I don't need to see any more pictures," among other statements, as being highly confident); NIJ, Eyewitness Identification*, A Trainer's Manual for Law Enforcement* (1999) ("Some witnesses will spontaneously include information about certainty (*e.g.*, 'That's him, I KNOW that's him,' or, 'It could be that one.'"). If the witness does not volunteer information about certainty, then the witness can be asked to state certainty in his/her own words.").
[141] NAS at 11 (emphasis added).

There is a common pattern among the DNA exoneration cases in general: the eyewitnesses whose high-confidence courtroom IDs contributed to wrongful convictions were not certain at the time of the initial identification. In fact, of the cases where the initial level of confidence could be ascertained, not one was made with high confidence.[142] The fact that, at that time, the effect of initial uncertainty was not properly understood should not be taken as an indictment of the reliability of an initial eyewitness identification.[143] "Had this simple fact been better understood by the legal system, many of the innocent defendants were convicted based, in part, on a high-confidence ID that occurred in court may never have been convicted in the first place."[144]

Archival and field studies demonstrate that highly confident witnesses are highly accurate.

In an archival analysis of files from the Sacramento City Police Department for crimes committed between 1987 and 1998, witnesses who made an identification in 56 live six-person lineups were asked to rate their confidence using what is essentially a two-point scale with verbal descriptors indicating either high confidence ("I am sure…") or low confidence ("I am not sure, but I think…").[145] The probability that a suspect was identified increased dramatically with confidence. Indeed, almost all of the filler ID errors were made with low confidence; for high-confidence IDs, 18 out of 19 (95%) were suspect IDs. Moreover, in cases where there was substantial extrinsic evidence of guilt, 12 out of 13 (92%) "sure" witnesses picked the suspect. *Id*. at 483.

Indeed, Wixted et al. (2015) concluded from the initial findings that "the relationship between confidence and accuracy in the real world mirrors the impressive relationship observed in experimentally controlled research in the sense that low-confidence IDs (the kind of IDs that have often contributed to wrongful convictions) are relatively error prone whereas high-confidence IDs are much less so."[146]

More recently, researchers analyzed 348 photo arrays that were blindly administered both simultaneously and sequentially in Houston. There was corroborating evidence in approximately 85% of the cases in which the suspect was picked, "suggesting that suspects identified by an eyewitness were more likely to be guilty than suspects who were not identified by an eyewitness."[147] This article concluded, "confidence in an eyewitness identification from a fair lineup is a highly reliable indicator of accuracy."[148]

The results of all identification procedures should be fully disclosed to the jury.

[142] Garrett, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG, (Harvard U. Press 2011).
[143] There is insufficient information on the Innocence Project's website to assess the initial confidence of the witnesses in all but a handful of cases. We understand that the Innocence Project accepts only a very small percentage of the cases presented to it and, of those, a significant percentage of DNA analyses confirm the defendant's culpability.
[144] Wixted & G. Wells (2017), *supra*, at 49.
[145] Behrman & Davey, *supra*, at 482.
[146] Wixted et al. (2015), *supra* at 521.
[147] Wixted, Mickes, Dunn, Clark & W. Wells, *Estimating the reliability of eyewitness identifications from police lineups,* Proceedings of the National Academy of Sciences, 304, 305 (2016).
[148] Wixted et al (2016), *supra*, at 304.

The NAS recommended that judges "make juries aware of prior identifications, the manner and time frame in which they were conducted, and the confidence level expressed by the eyewitness at the time."[149] As a general rule, such evidence would be introduced in the government's case-in-chief. Law enforcement agencies should require investigators to record details about the identification procedure and the results of the procedure, including the exact words spoken by each witness and any expression, gesture, or body language such as pointing, nodding, shaking one's head, or showing emotion during the identification procedure. Details of additional identification procedures, if any, also should be to be recorded. Of course, the jury would be aware of an identification made at trial. There should be no secrets about the strength of the identification from start to finish. Any increase or decrease in the witness's confidence can be assessed for what it is. The fact that studies in the laboratory show that a subject's confidence can grow over time adds nothing to the calculus in a given case where it would be obvious to the jury that a witness's confidence has or has not changed.

Conclusion

The research described above does not support the claim that "eyewitness confidence is generally an unreliable indicator of accuracy." This research, therefore, should not be the basis for any recommendations to that effect.

[149] NAS at 6, Recommendation #7.

SUMMARY
STRESS

> *The overall negative impact of heightened stress on accuracy of face identification was due entirely to a substantial effect on hit rate for TP [target present] lineups. The correct rejection rate for TA [target absent] lineups was unaffected by stress level.*[150]

This critical sentence in a well-regarded meta-analysis has been overlooked or minimized by those who have testified, advised the courts, or written about eyewitness identification. What it means is that stress does not increase the risk that an innocent person will be identified as the perpetrator; stress does increase the risk that a guilty person will not be identified as the perpetrator. As a consequence, the claim that stress reduces accuracy does not apply to witnesses who have identified the suspect and are likely to testify at trial.

Overall, the effect of stress on eyewitness identification is less uniform and definitive than it has been represented to be. Research presents a picture that is quite complex, indicating that stress can either help or hinder memory depending on a variety of factors. For some people, stress reduces the chance that they will be able to pick the perpetrator in a show up, lineup, or photo array; for others, stress increases the chance; and for the rest, there is no difference. For ethical reasons, laboratory researchers cannot duplicate real crimes. Thus, the generalizability of their research to the real world has been challenged. In addition:

• In the Deffenbacher meta-analysis, "one study was responsible for most of [the] difference in effect sizes." That study was not published in a peer-reviewed journal and involved only 54 subjects. Another similar study concluded that higher stress yields greater accuracy.

• In the Morgan study,[151] there were fewer false identifications of "innocent suspects" in the high stress condition than in the low stress condition.

• There is no uniform standard as to what constitutes a high level of stress and no way to measure it.

• Laboratory studies use stressors that are external to the event to be remembered and may serve as distractors.

• Memory is better for central details than peripheral ones.

Regardless of impact of stress on eyewitness identification generally, research now indicates that highly confident witnesses are highly accurate. A recent field study involving robberies, for example, which presumably involved stress, concluded that "confidence in an eyewitness identification from a fair lineup is a highly reliable indicator of accuracy." It appears, then, that stress does not adversely affect the accuracy of the suspect identifications that are made, although the suspect identification rate was fairly low.

---

[150] Deffenbacher, Bornstein, Penrod & McGorty, *A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory*, Law & Hum. Behav., 687, 695 (2004) (emphasis added).
[151] Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Baranoski & Southwick, *Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress*, Int'l J. Law & Psychiatry, 265 (2004).

STRESS

> *The overall negative impact of heightened stress on accuracy of face identification was due entirely to a substantial effect on hit rate for TP [target present] lineups. The correct rejection rate for TA [target absent] lineups was unaffected by stress level.*

Deffenbacher, Bornstein, Penrod & McGorty, *A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory,* Law & Hum. Behav., 687, 695 (2004) (emphasis added).

> *Morgan showed that stress reduced the correct identification rate in target present lineups, but had no effect on the mistaken identification rate in TA [target absent] lineups.*

Clark & Wells, *On the Diagnosticity of Multiple-Witness Identifications*, Law & Hum. Behav., 406, 415 (2008), *citing* Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Baranoski & Southwick, *Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress*, Int'l J. Law & Psychiatry, 265 (2004).

Research as Summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on stress as follows: "Even under the best viewing conditions, high levels of stress can reduce an eyewitness's ability to recall and make an accurate identification. Therefore, you should consider a witness's level of stress and whether that stress, if any, distracted the witness or made it harder for him or her to identify the perpetrator."[152]

Analysis

Stress does not increase false identifications.

Expert witnesses have repeatedly testified that stress reduces identification accuracy. This could leave the impression that an innocent person is more likely to be identified when the witness is under stress than when the witness is not under stress. However, the research does not support this impression. The studies upon which defense experts rely (including ones they have authored themselves) show that high stress decreases the identification of the correct "target," but it does not increase the identification of an innocent "target." Thus, it is not true that an innocent person is more likely to be chosen in a high- than low-stress circumstance, and it would be both inaccurate and misleading to instruct or testify otherwise.

The Special Master in *Henderson* (and many experts) relied primarily on the two articles cited above, Deffenbacher and Morgan, in concluding that high stress reduces identification accuracy.[153] However, the Deffenbacher meta-analysis found that "the overall negative impact of heightened stress on accuracy of face identification was due *entirely* to a substantial effect on hit rate for TP [target-present]

[152] Press Release, New Jersey Courts, *Supreme Court Releases Eyewitness Identification Criteria for Criminal Cases,* (July 19, 2012), http://www.judiciary.state.nj.us/pressrel/2012/pr120719a.htm; *see State v. Henderson*, 27 A.3d 872, 904 (2011).

[153] Report of the Special Master, *New Jersey v. Henderson*, 43 (6/18/2010) (citing transcripts and documents).

lineups. The correct rejection rate for TA [target-absent] lineups was unaffected by stress level."[154] In the Morgan study, there were fewer false identifications in the high stress condition than in the low stress condition. Because high stress does not increase the risk that an innocent suspect will be identified, the overall reduction in identification accuracy is irrelevant to the issue the jury must decide.

In the Morgan study, high stress reduced the percentage of false identifications.

The study conducted by Morgan et al. (2004) has been (and continues to be) widely cited in support of the proposition that stress reduces eyewitness identification accuracy.[155] However, while the results do in fact show that stress impairs memory, the implications of their findings have been widely misinterpreted. Table 1 shows that a high level of stress significantly reduced the identification of a "guilty suspect": 32% in high stress condition contrasted with 68% in the low-stress condition (aggregated across the three types of lineups they conducted). However, the same table shows that high levels of stress *reduced* the identification of an "innocent suspect": 39% in the high-stress condition contrasted with 46% in the low-stress condition.[156]

The critical point is that high stress dramatically lowered the correct identification rate (53%) while it also lowered the false identification rate by a lesser amount (15%). This result makes sense. Stress can be high enough that a witness will fail to form a memory of the perpetrator, making it less likely that the witness will be able to identify the guilty suspect from a later lineup. However, there is no reason to expect – and research does not support the proposition – that high levels of stress would implant a false memory of a particular innocent suspect who will later be placed in a lineup. It is therefore a mistake to suggest that the overall reduction of accuracy in this study means that high levels of stress increase the likelihood of a false identification. According to this influential (but misunderstood) study, high levels of stress have no such effect.

In the Deffenbacher meta-analysis, one study was responsible for most of the difference in effect sizes.

[154] Deffenbacher, Bornstein, Penrod & McGorty, *A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory,* Law & Hum. Behav., 687, 695 (2004). While the absence of an effect on target absent arrays was mentioned in the hearing before the Special Master, its significance was not highlighted.

[155] In this study, soldiers participated in SERE [Survival, Evasion, Resistance, and Escape] training, followed by sleep and food deprivation, and a high- and a low-stress interrogation. They were then asked to identify their interrogators. The high-stress interrogation was later described as follows: "Throughout the interrogation, the student is required to face the instructor and must maintain eye contact. In addition, the student must always adopt a height that is less than that of the instructor by bending or straightening his or her knees. Failure to comply with this rule results in physical punishment to the student by the interrogator. Thus, students must be attentive to the face and relative height of the instructor . . . . The various types of physical confrontation have been made public . . . and include facial slaps, abdominal punches, walling (slamming the student into the wall) and stress positions." Morgan, Southwick, Steffian, Hazlett & Loftus, *Misinformation can influence memory for recently experienced, highly stressful events,* Internat'l J. Law & Psychiatry, 11, 12-13 (2013). Most witnesses, even victims, are not subjected to this kind of extreme physical abuse.

[156] This data was not presented directly in Table 1, nor was it discussed in text. But it can be calculated easily by subtracting the correct rejection rate in each condition from 100%. The false identification rates likely are considerably lower, but because the lineups differed in size and the number of "suspects" in each, the normal method of calculating them by dividing by the number in the array cannot be used.

The Deffenbacher meta-analysis assessed 27 earlier studies involving high- and low-stress conditions, five of which involved facial recognition studies and 22 of which involved "the more ecologically valid eyewitness identification tradition."[157] In the facial recognition studies, there was virtually no difference in the mean proportion of correct identifications in the high- (.56) and low- (.58) stress conditions. In eyewitness identification studies, there was a difference in the mean proportion of *correct* identifications in the high- (.37) and low- (.53) stress conditions. However, "[o]ne study [by Buckhout] was responsible for most of this difference in effect sizes."[158] The authors minimize this fact by claiming that the study "was a rather realistic, *live* staged crime, rather than a filmed one."[159] Even if it were true that the staged crime was realistic and representative, the Buckhout study was not published in peer-reviewed journal and involved only 54 students. Another set of researchers who conducted a similar live staged crime experiment found that, "to the extent that there was a correlation ($p<.10$), it was in the direction of higher arousal being associated with more accurate identification."[160] At best, then, this research is in equipoise.

As to the remaining 21 studies in the Deffenbacher meta-analysis where, apparently, little or no effect was found, "[i]t is astonishingly naïve of psychologists to believe that most undergraduates would find the content of such films [with a violent ending] stressful. Most Americans see so many shootings and other violence in the movies and on television that they become inured to filmed violence . . . Studies that employ such films as stimuli generally bear no relationship to the impact of real-life violence or threatened violence on victims of crime."[161] Discounting the lack of an effect in 21 out of 22 "more ecologically valid" studies undermines the strength of Deffenbacher's conclusions and may even suggest bias.[162]

---

[157] Deffenbacher et al. (2004), *supra,* at 697. Facial recognition studies usually involve looking at a number of faces and then trying to identify them in a larger or different set of faces.

[158] Deffenbacher et al. (2004), *supra,* at 697.

[159] This crime involved a purse snatching in a classroom. The thief, who said he had forgotten a book, "grabbed a confederate's purse and fled out the door. The confederate screamed as the assailant ran out of the class. The instructor . . . ran after the assailant. A few minutes later, the instructor . . . returned and informed the students of the experimental nature of the incidents . . . ." Buckhout, Alper, Chern, Silverberg & Slomovits, *Determinants of eyewitness performance on a lineup,* Bull. Psychon. Soc., 191, 191 (1974). The correct identification rate was only 13.5%, much lower than in most laboratory and real-life studies. Query whether the students' attention was drawn to the scream or the fleeing "thief." The whole incident lasted 21 seconds. The authors do not reveal the length of time the thief's face was visible as he fled. Moreover, the "thief" in the target absent array was "a close look-alike" to the target, something that cannot happen in real life except by chance.

[160] Egeth, *Emotion and the Eyewitness,* in THE HEART'S EYE: EMOTIONAL INFLUENCES IN PERCEPTION AND ATTENTION, 245, 252, (Niedenthal & Kitayama, eds. 1994), *citing* Hotch & Bothwell, *Arousal, description and identification accuracy of victims and bystanders*, Journal of Social Behavior & Personality, 481 (1990)). Stress is often referred to in the literature as "arousal."

[161] Yuille & Daylen, *The Impact of Traumatic Events on Eyewitness Memory*, in EYEWITNESS MEMORY: THEORETICAL AND APPLIED PERSPECTIVES, 155, 157-158 (Thompson et al., eds. 1998).

[162] Citing the Deffenbacher meta-analysis, among many others, the National Academy of Sciences [NAS] commented that "none of the reviews met all current standards for conducting and reporting systematic reviews, and few met even a majority of these standards, making assessment of the credibility of their findings problematic. After examining the reviews, the committee concluded that the findings may be subject to unintended biases and that the conclusions are less credible than was hoped." National Academy of Sciences, Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 75-76 and n.9 (2014) (footnote omitted).

In an earlier review "by Deffenbacher (1983) across 21 studies on arousal and eyewitness memory, 10 studies showed that high arousal levels increased eyewitness accuracy, whereas 11 studies showed a lower accuracy at high arousal levels. Deffenbacher resolved this inconsistency in research findings by reclassifying the arousal levels in the various studies and then concluding that most of the studies supported the inverted U-form relationship between arousal and memory, known as the Yerkes-Dodson (1908) law."[163] However, this conclusion has been criticized on the ground that "the Yerkes-Dodson law does not constitute an appropriate description of the relationship between emotional stress and memory performance; thus, this theory is not a pragmatically useful abstraction in evaluating eyewitness situations."[164] "It is interesting to note that the 10 studies showing either improved performance or no effect of stress typically involved staged live events and the studies showing a negative effect of stress did not involve live events."[165] The Buckhout study, therefore, appears to be an outlier that has had a disproportionate effect on the assessment of stress.

"[I]t is clear that emotional events are indeed remembered differently than neutral or ordinary events. However, the pattern of results from various studies proves that there are no real grounds for a simple relationship between intense emotion and memory, that is, that the more negative the emotion or stress, the poorer the memory or the opposite, that intense emotion leads to generally detailed, accurate and persistent memory. In fact, recent research in this field . . . shows that the way emotion and memory interact is a very complex matter."[166]

Stress certainly can be sufficiently high that memory of a perpetrator will be impaired. Indeed, stress may have been sufficiently high in the Morgan et al. (2004) and in the Buckhout study, which reported similar effects of stress on the correct ID rate. The key point is that none of the available evidence suggests that any such reduction in the correct ID rate implies that there is a corresponding increase in the false ID rate. Instead, the available evidence suggests that stress does not increase the risk that an innocent suspect will be chosen.

There is no uniform standard as to what constitutes "a high level of stress" and no way to measure it.

The Court in *Henderson* acknowledged the lack of a uniform standard for high stress, stating "[t]here is no precise measure for what constitutes high stress, which must be assessed based on the facts presented in individual cases."[167] Research does not provide much guidance. As one of the experts in the *Henderson* case acknowledged later:

> Over the years, we have not, as a field, been able to get a good handle on the effects of stress, partially because it's hard to perform the experimental conditions that would

[163] Christianson, *Emotional Stress and Eyewitness Memory: A Critical Review,* Psychol. Bull., 284, 289, (1992) (citing Deffenbacher, *The influence of arousal on reliability of testimony*, in EVALUATING WITNESS EVIDENCE, 235 (Wiley 1983)). The Yerkes-Dodson law "states that performance improves with increases in arousal [i.e., stress] up to some optimal point and then declines with further increases," Deffenbacher et al. (2004), *supra*, at 688, hence, the "inverted U-form." Of course, even if Yerkes-Dodson were a valid construct, a proposition with which we do not agree, assessing where a witness falls on the continuum would be difficult. If he were on the upside, then stress would increase accuracy; if on the downside, then it would decrease it.

[164] Christianson (1992), *supra,* at 303.

[165] Yuille & Cooper, *Challenging the Eyewitness Expert,* in Ziskin & Faust (Eds.), COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY, 658, 689 (6th edition 2012).

[166] Christianson (1992), *supra,* at 302-303, (citations omitted) (Y-D is not the correct paradigm).

[167] *State v. Henderson*, *supra*, 27 A.3d at 904.

> lead to high levels of stress and be within ethical boundaries. So there hasn't been very much on that.

Testimony of Roy Malpass, Ph.D., in *United States v. Thomas*, No. 1:13-CR-03897-MV, at 109 (D.N.M. 1/22/15).[168]

The kind of stress in laboratory research does not reflect the emotional response a witness might experience when seeing or being the victim of a crime. For example, some experiments look at effects of stress by asking some subjects to remember details from emotional events and other subjects to remember details from neutral events.[169] Others attempt to simulate stress by having subjects watch a video of a crime or emotional event.[170] Still others use noise, mild electric shock, the use of carbon dioxide, or other distractors.

Several reviews of the literature have criticized the research on stress and questioned its generalizability to real eyewitnesses.[171]

Laboratory experiments often use stressors that are external to the event and may serve as distractors.

Because they cannot replicate crimes in their studies, laboratory researchers may use stressors that are external to the event to be remembered and may operate as distractors.[172] "[E]xtraneous sources of arousal like loud noise, failure stress, worry about a threatening experimental situation, and so on are likely to make a person nervous and emotional and will accordingly distract him or her from the TBR [to-be-remembered] information. If the emotionally arousing agent is related to the TBR event and if one is not distracted by an extraneous source of arousal that is stronger in intensity than the TBR event, there is no evidence that high arousal impairs memory performance. . . ."[173] "There is reason to believe that arousal extrinsic to the to-be-remembered materials may influence memory in a way different from emotion somehow related to the to-be-remembered material. That is 'memory while emotional' may be distinct from 'memory for emotional materials.'"[174] Indeed, "studies that use a stressor that is external

[168] *See* Herve, Cooper & Yuille, *Biopsychosocial Perspectives on Memory Variability in Eyewitnesses* in APPLIED ISSUES IN INVESTIGATIVE INTERVIEWING, EYEWITNESS MEMORY AND CREDIBILITY ASSESSMENT, 99, 106 (Springer 2013) ("Laboratory-based methodologies are, for ethical reasons, unable to evoke remarkable memories as the stimuli used cannot produce extreme stress or trauma.").

[169] *See, e.g.*, Heuer & Reisberg, *Vivid Memories of Emotional Events: The Accuracy of Remembered Minutiae*, 18 Mem. & Cog., 496, 498-99 (1990) (subjects watched a slide sequence with either a neutral or more violent ending and given a memory test two weeks later; those in the arousal group did better on recall of both central and peripheral details); Houston, Clifford, Phillips & Memon, *The Emotional Eyewitness: The Effects of Emotion on Specific Aspects of Eyewitness Recall and Recognition Performance*, Emotion, 118, 119 (2013).

[170] *See*, *e.g.*, Echterhoff & Wolf, *The Stressed Eyewitness: The Interaction of Thematic Arousal and Post-Event Stress in Memory for Central and Peripheral Event Information*, Frontiers in Integrative Neuroscience (August 23, 2012) (Subjects watched a video depicting a burglary under high or low thematic arousal); Houston et al. (2013), *supra,* at 120 (participants viewed either a mugging or conversation on video).

[171] Christianson (1992), *supra,* at 284; Egeth (1994), *supra,* at 252; Yuille & Cooper (2012), *supra; See* Wells, Memon & Penrod, *Eyewitness Evidence: Improving its Probative Value*, Psychological Science in the Public Interest, 46, 49 (2006) ("There may be limits to generalizing from experiments to actual cases.").

[172] *See* Yuille & Cooper (2012), *supra,* at 688.

[173] Christianson (1992), *supra,* at 297.

[174] Burke, Heuer & Reisberg, *Remembering emotional events*, Memory & Cognition, 277, 288 (1992).

to the eyewitness event are actually studies of unpleasant distracters rather than studies of the effects of witnessing or experiencing a stressful even such as a criminal act."[175]

Memory is better for central details than peripheral ones.

Although they do not necessarily involve eyewitness identification (*i.e.*, these studies have examined eyewitness memory in the broader sense), a number of studies have concluded that stress actually improves memory for spatially and temporally central details of an event while it may undermine memory for peripheral ones:

- "One can see that highly negative emotional events are relatively well retained, both with respect to the emotional event itself and with respect to the *central, critical detail* information of the emotion-eliciting event – the information that elicits the emotional reaction."[176]
- "Arousal improved memory for gist, for [basic level visual information], and for details that happened to be associated spatially with the event's center. In contrast, arousal undermined memory for background detail."[177]
- "The emotional memory enhancement appears to be especially pronounced for central aspects of the arousing item, whereas emotional arousal often impairs memory for peripheral details. This has been interpreted as a result of attentional narrowing."[178]
- "Stress and arousal . . . interact to enhance memory for aversive aspects of a negative event such that this information is subsequently remembered with accuracy and [is] less vulnerable to the incorporation of misinformation."[179]
- "When science is interpreted properly, the evidence shows that traumatic events – those experienced as overwhelmingly terrifying at the time of their occurrence – are highly memorable and seldom forgotten."[180]
- "Arousal is likely to be associated with challenging, important or threatening events, for which fast and focused responding is critical. Enhanced processing of salient, surprising or goal-relevant stimuli should improve performance under such circumstances. Later, remembering the high priority information from the event could improve future strategies for dealing with similar situations. But the increased advantage of high

[175] Yuille & Daylen (1998), *supra,* at 157.
[176] Yuille & Cooper (2012), *supra* (citing other studies).
[177] Burke et al. (1992), *supra,* at 289. Finding that "[t]he present data fit with other findings alleging a narrowing (or redirecting) of attention during emotional events," these authors speculated that there might be a difference between material central to the plot and that central to the viewer's attention, and called for more research on this distinction. *Id.* at 289.
[178] Echterhoff & Wolf (2012), *supra,* at 1 (citations omitted). In this study, the mean overall accuracy rates for central details varied only between .78 and .86 in four conditions, those at the highest stress level being the most accurate. *Id*. at 7, Table 1.
[179] Hoscheidt, LaBar, Ryan, Jacobs & Nadel, *Encoding negative events under stress: High subjective arousal is related to accurate emotional memory despite misinformation exposure,* Neurobiology of Learning & Memory, 1, 10 (In press 2013).
[180] McNally, *Debunking Myths About Trauma and Memory*, Can. J. Psychiatry, 815, 821 (2005) ("Release of stress hormones during aversive or traumatic events strengthen memory for the traumatizing experience. Intense arousal enhances memory for the core features of the arousing event; it does not attenuate it.").

> priority information comes at the expense of low priority information that garners even fewer neural resources under arousal than it would otherwise."[181]

One study involved a film with a violent or neutral ending. In the violent ending, a boy was shot in the face. The researchers deemed the number on the boy's jersey "critical" although, in real life, it would probably be inconsequential. Recall for that detail was much smaller for those who saw the violent ending (4%) than the neutral ending (28%), but overall the average percentage of accurate recall for other items was 80% for the violent ending compared to 84% for the neutral ending, "a small difference,"[182] and one that might be attributable to the subject's focus of attention.

The same event affects different people differently.

The assumption that crime creates high stress in all victims and witnesses also is unwarranted. The London Dungeon study is frequently cited by experts for the proposition that high stress reduces identification accuracy.[183] Visitors volunteered for a study whose purpose was not disclosed. While they were in the Horror Labyrinth, an actor, wearing very pale facial makeup with wounds or scars, temporarily blocked the path of the visitors. Forty-five minutes after they left the Horror Labyrinth, the visitors were given an anxiety questionnaire, asked to provide a description of the actor, and shown a nine-photo simultaneous target-present array to identify that actor. Based on responses to the anxiety questionnaire, the group of visitors was divided down the middle: the half that was above the median ("high anxiety") made fewer accurate identifications of the actor in the array (29%) than the half that was below the median ("low anxiety") (81%).[184] Thus, the identical stressor caused different subjects different levels of stress.

Recognizing that Deffenbacher found no stress effect on mistaken identification in target absent arrays, the London Dungeon study included only target present arrays.[185] This study, then, does not address the only issue of concern to a jury, that is, the effect of high stress on mistaken identifications.

As significantly, although the authors do not provide details, they found that for participants who made an identification – whether high or low stress – the "accuracy of identification was reliably associated with confidence."[186]

---

[181] Mather & Sutherland, *Arousal-biased competition in perception and memory*, NIH Public Access, Perspect. Psychol. Sci., 1, 19 (March 24, 2011). Some researchers question whether the face of the target is a central detail. One study found, for example, that "while emotional participants again provide a more complete description of the perpetrator, they are less able [by about 12%] than their neutral counterparts to recognize the perpetrator from a photographic lineup." Houston et al. (2013), *supra,* at 118.

[182] Egeth (1994), *supra,* at 251, *citing* Loftus & Burns, *Mental shock can cause retrograde amnesia*, Memory & Cognition, 318-323 (1982). Retrograde amnesia is loss of memory-access to events that occurred, or information that was learned, before an injury or the onset of a disease.

[183] Valentine & Mesout**,** *Eyewitness Identification Under Stress in the London Dungeon*, Appl. Cog. Psychol., 151 (2008). The London Dungeon is a cross between a haunted house and a wax museum. The "Horror Labyrinth, a maze of floor to ceiling mirror walls set among Gothic vaults, was "[d]esigned to disorient visitors, it is dark, crowded and there was a soundtrack playing the rhythm of a heart-beat and various scary noises and screams." *Id*. at 154. There are a number of "scares" in the Horror Labyrinth in addition to the actor described in text.

[184] Valentine & Mesout (2008), *supra,* at 158, Table 3. In real life, it would be inappropriate, and probably impossible, to administer anxiety questionnaires to victims and witnesses.

[185] Valentine & Mesout (2008), *supra,* at 153 ("previous research has shown an effect of stress on eyewitness identification on target present but not on mistaken identification from target-absent lineups").

Some people do better under high stress; stress does not increase – and may in fact reduce – the chance that an innocent suspect will be chosen.

"Empirical data yields a picture that is quite complex, indicating that emotional arousal can either aid or hinder memory depending on a variety of factors."[187]

In the Morgan study, discussed above, with respect to the identification of "guilty suspects," "42-45% [about 43%] of subjects performed equally well or poorly across the stress conditions, 42-50% [about 46%] of subjects performed better in the low-stress condition, and . . . 8-13% [about 11%] of subjects performed better in the high-stress condition."[188] Thus, stress had no effect on or actually improved identification accuracy for more than half of the subjects.

"This kind of variability is what one observes when working with victims, witnesses, and offenders in the criminal justice system. There is no typical or average way of responding to violence, threats of violence, sexual assault, hostage taking, etc. Instead, there is a range of the impact of stress all the way from a completely debilitating effect on memory to improving memory."[189]

It should not be said, therefore, that "high levels of stress can reduce an eyewitness's ability to recall and make an accurate identification," without also saying that high levels of stress can increase an eyewitness's ability to make an accurate identification, or can have no effect at all. One should also add that high stress does not increase false identifications.

Highly confident witnesses are highly accurate.

Assuming for the sake of argument that stress reduces the proportion of witnesses who correctly identify the target, what difference does it make? In real life, people who select a filler/foil or who do not select anyone are not identification witnesses at trial.[190] The underlying flaw in these studies is that a reduction in the overall proportion of correct responses has some bearing on the accuracy of the subset of witnesses who pick the target (or, in real life, the suspect) even if they were highly stressed when they witnessed the event/crime.[191] This leap of logic has no basis. High stress at the time of exposure may reduce the proportion of witnesses who are able to make an identification (so overall accuracy declines), but it does not reduce the accuracy of the subset of witnesses who nevertheless

---

[186] Valentine & Mesout (2008), *supra,* at 158.

[187] Bornstein, Bryant & Zickafoose, *Intuitions about Arousal and Eyewitness Memory: Effects on Mock Jurors' Judgments*, Law & Psychology Review, 109, 111 (Spring 2008). *Accord* Christianson (1992), *supra*, at 303; Yuille & Daylen (1998), *supra,* at 160.

[188] Morgan et al. (2004), *supra,* at 273.

[189] Yuille, *The Challenge for Forensic Memory Research: Methodolotry*, in APPLIED ISSUES IN INVESTIGATIVE INTERVIEWING, EYEWITNESS MEMORY AND CREDIBILITY ASSESSMENT, 1, 8 (Springer 2013). *See also* Herve et al. (2013), *supra,* at 27 ("[T]he results [in Morgan] confirm that stress/arousal has complex effects on eyewitness memory.").

[190] In real lineups, suspects are picked around 40% of the time. *See* Slater, *Identification Parades: A scientific evaluation,* Police Research Group, Home Office (1994) (36%); Wright & McDaid, *Comparing System and Estimator Variables Using Data from Real Lineups,* Appl. Cog. Psychol., 75, 77 (1996) (39.1%.); Valentine, Pickering & Darling, *Characteristics of Eyewitness Identification that Predict the Outcome of Real Lineups*, Appl. Cog. Psychol., 969, 970 (2003) (41%).

[191] As discussed above, stress appears to have no effect on target-absent arrays. Clearly, laboratory subjects who pick the target in a target-present array are 100% accurate.

make an ID with a particular level of confidence.[192] More specifically, as a general rule, stress at the time of exposure does not cause eyewitnesses to overestimate their confidence when they pick the suspect/target.

In a recent study of victims and witnesses to real robberies in Houston, published in the Proceedings of the National Academy of Sciences, the authors concluded that "confidence in an eyewitness identification from a fair lineup is a highly reliable indicator of accuracy."[193] Although this study did not examine stress specifically, it is significant because all of the cases involved robberies where, one can assume, at least some stress was present. Yet, highly confident witnesses were highly accurate.

Conclusion

The research shows that stress does not increase the chances that an innocent suspect will be chosen. To the extent that research is relevant at all, experience confirms that stress increases the capacity of some victims and witnesses to perceive and remember the perpetrator accurately, decreases the capacity of others, and does not change the capacity of the remainder. Different witnesses view and remember events differently. Caution should be used in generalizing from controlled research studies to real-world contexts, as stress in eyewitnesses to specific crimes cannot be replicated by laboratory experiments. Recommendations should not be based on inadequate, inconsistent, and irrelevant studies.

[192] *See* Confidence and Accuracy.

[193] Wixted, Mickes, Dunn, Clark & W. Wells, *Estimating the reliability of eyewitness identifications from police lineups,* Proceedings of the National Academy of Sciences, 304, 309 (2016).

SUMMARY

WEAPON FOCUS EFFECT

> *"[T]he crucial question appears to be what effect the presence of a weapon has on TA [target absent] lineups. Unfortunately, most studies did either not use TA lineups or simply reported results averaged across both TP [target present] and TA lineups ("correct decisions"). . . . [N]o conclusions can be drawn regarding the [weapon focus effect] on TA lineups and future research is clearly needed to remedy this deficit. This also implies that there seems not to be sufficient evidence for experts testifying for the defense on the weapon focus effect with respect to identification decisions, as this type of expert testimony typically focuses on factors contributing to the likelihood of false identifications."*

Kocab & Sporer, *The Weapon Focus Effect for Person Identifications and Descriptions: A Meta-analysis,* ADVANCES IN PSYCHOLOGY AND LAW, Vol. I, 71, 105 (Miller & Bornstein eds. 2016).

There are three fundamental problems with the current understanding of the weapon focus effect:

(1) As indicated by the quote above, there is insufficient evidence that the presence of a weapon increases false identifications as opposed to reducing true ones. This means that the weapon focus affect is unlikely to apply to witnesses who identify the suspect and testify at trial.
(2) The weapon focus effect has not been found in most real world studies – and when it has been found, it reduces the likelihood that the suspect (whether guilty or innocent) will be identified.
(3) There is virtually no difference in accuracy for witnesses/subjects who identify suspects/ targets with high confidence, that is, 95-100% accuracy for those who say they are 90-100% confident, whether a weapon is present or not.

Some other shortcomings in the research that have contributed to the misimpression include:

- It assumes that witnesses are distracted by a weapon and do not focus on the face; this may be true (and likely contributes to a lower identification rate), but real witnesses disclose to the police what they were looking at, thereby indicating whether they should be shown a photo array.
- Laboratory research cannot replicate real world crimes committed with weapons and, therefore, haS limited ecological validity.
- Researchers sometimes conflate feature accuracy with identification accuracy, thereby greatly inflating the weapon focus effect on the latter.
- In the laboratory, the weapon focus effect on identification accuracy is about 10%, described by researchers as "small," or "not of great magnitude." It disappears with longer (or very short) exposures, but there has been no systematic exploration of exposure duration in the literature.
- The weapon focus effect does not appear to occur if the witness/subject sees the perpetrator before s/he sees the weapon or if the perpetrator is close to the witness/subject.

- Until recently, the statistical methods used in WFE research did not provide meaningful guidance.

WEAPON FOCUS EFFECT

> *"[T]he crucial question appears to be what effect the presence of a weapon has on TA [target absent] lineups. Unfortunately, most studies did either not use TA lineups or simply reported results averaged across both TP [target present] and TA lineups ('correct decisions'). There were only four hypothesis tests of studies that reported false identifications in TA lineups as a function of manipulated WFE [weapon focus effect]. The results were completely contradictory, with two tests reporting an increase and two a decrease in false IDs, with odds ratios ranging for 0.40 to 2.40, making the calculation of an average effect size not only meaningless but statistically inappropriate. . . . [N]o conclusions can be drawn regarding the WFE on TA lineups and future research is clearly needed to remedy this deficit. This also implies that there seems not to be sufficient evidence for experts testifying for the defense on the weapon focus effect with respect to identification decisions, as this type of expert testimony typically focuses on factors contributing to the likelihood of false identifications."*

Kocab & Sporer, *The Weapon Focus Effect for Person Identifications and Descriptions: A Meta-analysis,* ADVANCES IN PSYCHOLOGY AND LAW, Vol. I, 71, 105 (Miller & Bornstein eds. 2016).

Research as Summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on the weapon focus effect as follows: "You should consider whether the witness saw a weapon during the incident and the duration of the crime. The presence of a weapon can distract the witness and take the witness's attention away from the perpetrator's face. As a result, the presence of a visible weapon may reduce the reliability of a subsequent identification if the crime is of short duration. In considering this fact, you should take into account the duration of the crime because the longer the event, the more time the witness may have to adapt to the presence of the weapon and focus on other details.[194]

Analysis

There is insufficient evidence that the presence of a weapon increases false identifications for expert testimony or jury instructions.

As the quote above indicates, laboratory research does not indicate that the presence of a weapon increases the risk of falsely identifying an innocent person as the perpetrator. Until more data are available that show such an effect – and that it is applicable to real world crimes and to witnesses who are most likely to testify at trial – there is no basis for recommendations on the weapon focus effect.[195]

[194] Press Release, New Jersey Courts, *Supreme Court Releases Eyewitness Identification Criteria for Criminal Cases,* (July 19, 2012), http://www.judiciary.state.nj.us/pressrel/2012/pr120719a.htm; *see State v. Henderson*, 27 A.3d 872, 905 (2011).

[195] One recent study found an increase in false identifications when a weapon was visible but, computed in terms of false suspect ID rates (using the standard method of dividing all target-absent IDs by lineup size), the difference was only 2% (12% v. 14%), and the authors did not scale these responses by confidence. Carlson, Dias,

Other shortcomings in the laboratory research on the weapon focus effect, discussed below, cast further doubt on this "estimator variable."

If the presence of a weapon has any effect in the real world, it reduces both true and false identifications.

A 2013 meta-analysis found that, after decades of research, "neither field nor archival studies have reported an effect of weapon presence on suspect identification or description accuracy."[196] In fact, real-world studies suggest that the presence of a weapon may *decrease* misidentifications (non-perpetrators in the lineup)[197] and enhance detail in eyewitness accounts.[198] Of course, it is more difficult and expensive to conduct field or archival research. But relying on studies that fall far short, even in assessing a weapon focus effect in the laboratory, is unwarranted.[199] One expert explained it this way:

Weatherford & Carlson, *An Investigation of the Weapon Focus Effect and the Confidence-Accuracy Relationships for Eyewitness Identification*, J. Appl. Res. Mem. & Cog., 1, 5, Table 1 (2016). As discussed below, when confidence was taken into consideration, a weapon had no negative effects.

[196] *See* Fawcett, Russell, Peace & Christie, *Of Guns and Geese: A Meta-Analytic Review of the 'Weapon Focus' Literature*, 19 Psychol., Crime & L., 35, 43 (2013); *id.* at 56 ("it has only once demonstrated even a marginally negative effect on actual suspect identification . . . and has even been shown to decrease false alarms in another preliminary analysis.") (citing Tollestrup, Turtle & Yuille, *Actual Victims and Witnesses to Robbery and Fraud: An Archival Analysis*, in ADULT EYEWITNESS TESTIMONY: CURRENT TRENDS AND DEVELOPMENTS, 144-160 (David Frank Ross et al., 1994), and Valentine, Pickering & Darling, *Characteristics of Eyewitness Identification that Predict the Outcome of Real Lineups*, 17 Appl. Cog. Psychol., 969 (2003)).

[197] *See, e.g.*, Valentine et al. (2003), *supra,* at 980 (an archival study of 584 witnesses and 295 lineups found the presence of a weapon had "no effect on the likelihood of identifying a suspect, but . . . [w]itnesses were *more likely* to identify a foil if a weapon was *not* present.") (emphasis added).

[198] *See* Yuille, Ternes & Cooper, *Expert Testimony on Laboratory Witnesses*, 10 J. Forensic Psychol. Practice, 238, 243 (2010) ("[T]he few [field] studies that have examined the impact of the presence of a weapon on actual eyewitnesses have provided little to no empirical support for the weapons focus effect . . . Rather . . . the presence of a weapon in actual criminal cases tends to either have no detrimental effect on eyewitness memory or enhances detail in eyewitness accounts.").

[199] *See, e.g.*, Carlson & Carlson (2014), *supra,* at 45 ("We caution the reader that this weapon effect could be driven by particular qualities of the mock crime video we used."); Fawcett et al. (2013), *supra,* at 62 ("Upon collating the data for analysis we repeatedly came across areas in need of further characterization."); DeCarlo, *A Study Comparing the Eyewitness Accuracy of Police Officers and Citizens*, at 65-66, (2010) ("[A]lthough video presentation is not a perfect way to measure the weapon focus effect, it has proven to yield acceptable results to the degree necessary for the current *study*.") (emphasis added); Pickel, Narter, Jameson & Lenhardt, *The Weapon Focus Effect in Child Eyewitnesses*, 14 Psychol., Crime & L., 61, 70 (2008) ("The experiences of participants in the present study differed in several ways from the experiences of actual witnesses."); Cooper, Kennedy, Herve & Yuille, *Weapon Focus in Sexual Assault Memories of Prostitutes,* 25 Int'l J. Law & Psychiatry, 181, 182 (2002) ("It is quite possible that the concept of weapon focus is a laboratory phenomenon and is thus not applicable to actual criminal events."); Pickel, *The Influence of Context on the "Weapon Focus" Effect,* 23 Law & Hum. Behav., 299, 310 (1999) ("One limitation of the current study is that the stimuli were videotaped rather than live events, and the participants knew that they were not watching real crimes in progress. *Almost all existing studies on weapon focus share this limitation."*) (emphasis added); Mitchell, Livosky & Mather, *The Weapon Focus Effect Revisited: The Role of Novelty*, Legal & Criminol. Psychol., 287, 291 n.1, 300 (1998) ("Albeit the criticism against the ecological validity of using videotaped scenarios as stimuli in weapon focus studies is acknowledged and well taken . . . we chose to use such stimuli so that the present results could be compared more easily to previous studies . . . . It is important to note that the present study is itself of limited ecological validity, and so it may be ill-advised to try to map these results onto real-world crime situations too closely."); Steblay, *A Meta–Analytic Review of the Weapon Focus Effect,* 16 Law & Hum. Behav., 413, 422 (1992) ("It may be argued that real-life crime events include so many

"It is possible that participants in traditional research studies engage different strategies and behavioral patterns than do eye-witnesses to criminal events . . . . It is possible that weapons make criminals feel less vulnerable and therefore act less cautiously, spend more time at the crime scene, or venture closer to eyewitnesses thus diminishing the weapon focus effect . . . . Alternately, it could be that actual crimes are too complex, with too many other influential factors for the presence of a weapon to have a significant impact on eyewitness memory."[200]

The most recent (2016) police department field study found evidence consistent with what is observed in laboratory studies: "In this sample . . . eyewitnesses who were exposed to a weapon . . . made fewer suspect identifications and more rejections than witnesses who were not exposed to a weapon . . . ."[201] Consistent with laboratory studies, the presence of a weapon makes it less likely that a witness will form a memory needed to identify a guilty suspect (decreasing the correct ID rate in laboratory studies and decreasing the overall suspect ID rate in the field) without having the effect of somehow implanting a false memory of an innocent suspect who later appears in a lineup.

Highly confident witnesses are highly accurate even when a weapon is clearly visible.

A recent (2016) article makes clear that for highly confident witnesses – those who are most likely to testify at trial – "weapon presence, whether visible or concealed, [does] not negatively impact the confidence-accuracy relationship. In fact, participants were best calibrated when the weapon was clearly visible," and all conditions produced under-confidence not over-confidence.[202] Significantly, using the confidence-accuracy characteristic [CAC], which plots the accuracy of suspect identifications (whether "guilty" or "innocent") at each level of confidence, the authors found that there was virtually no difference in accuracy for highly confident witnesses in the three conditions (no weapon, weapon concealed and weapon shown):

> Essentially CAC curves simplify evaluation of the [confidence-accuracy] relationship for triers of fact, as they can simply depict low versus medium versus high confidence . . . . Rather than drawing attention to levels of over- versus under-confidence across the entire confidence range, as calibration curves allow, and which would be important for theoretical evaluations of the CA [confidence-accuracy] relationship, these CAC curves focus on the most relevant range of the confidence scale, namely the eyewitnesses who exhibit the highest confidence. These curves address the question: can identifications made by highly confident eyewitnesses (those most likely to make it to trial) be trusted? In other words, are these identifications highly accurate? [This data] shows that they are.[203]

---

stimuli that the hypothesized weapon focus effect becomes irrelevant or insignificant in magnitude. The problem of ecological validity of these laboratory-based data can only be addressed here by noting that those researchers who have attempted to maximize complexity by adding *additional bystanders* or *noise* . . . has not effectively eliminated the effect.") (citations omitted and emphasis added).

[200] Fawcett et al. (2013), *supra,* at 19 (citations omitted).

[201] W. Wells, Campbell, Li & Swindle, *The characteristics and results of eyewitness identification procedures conducted during robbery investigations in Houston TX*, Policing, An Internat'l J. Police Strategies & Management, 601, 613-614, Table II (2016). The data for this study were collected in 2013.

[202] Carlson et al. (2016), *supra*, at 1, 6.

[203] Carlson et al. (2016), *supra*, at 7; *id*. at 8 ("Relatively high confidence after choosing from a lineup tended to indicate high accuracy in that decision (Figs. 2 and 3), regardless of a weapon being shown, concealed, or absent").

In all three conditions the accuracy rate was 90% or greater for subjects who said they were 90-100% confident.[204] As in studies of other effects, there are fewer highly confident witnesses when a weapon was shown, but the accuracy of highly confident witnesses is not diminished because of it.[205] Thus, the science does not support the proposition that "the presence of a visible weapon may reduce the reliability of a subsequent identification if the crime is of short duration" for those witnesses who are likely to testify at trial. It would be both incorrect and misleading to suggest otherwise.

If a witness was not distracted by the presence of a weapon, then there could be no "weapon focus effect."

The proposed instruction is based on the assumption that when a weapon is present, a witness will focus on the weapon and not on the face or features of the perpetrator – having been distracted by stress or novelty of the weapon. If, however, the presence of a weapon does not distract the witness, in whole or part, and does not take attention away from the perpetrator's face, then the witness's ability to identify the perpetrator would not be diminished and the reliability of the witness's identification would not be impaired.

The majority of the laboratory results and the results of a recent police department field study are consistent with the idea that a weapon may make eyewitnesses less likely to choose anyone from a lineup—thus lowering overall "identification accuracy."[206] This says little about the real eyewitnesses to real crimes who do identify the suspect, particularly with high confidence.

Laboratory research does not replicate real-life conditions or circumstances.

As researchers themselves acknowledge, there are many constraints in the laboratory that limit the applicability of their findings to the real world. Among them are the restrictions on human subject research, the failure to ask the subjects what they were looking at, the use of a "forced choice" format, and the absence of confidence statements to measure the strength of an identification. In addition, many of the laboratory studies on weapon focus address "feature accuracy" which is different from "identification accuracy" and should not be conflated with it.

---

[204] Carlson et al. (2016), *supra*, at 7, Figure 3. "There was no significant difference in the number of correct IDs between weapon concealed and no weapon conditions." *Id*. at 5. There was no difference accuracy between the weapon shown and weapon concealed conditions and slightly higher accuracy for the no weapon condition – all in the 85-90% range – for people who assessed their confidence at 70-80%. For lower confidence witnesses (0-60%), accuracy dropped off more steeply for the weapon shown condition than the no weapon or weapon concealed conditions. *Id.* at 7, Figure 3. Ninety-four percent of the participants in the concealed weapon condition reported that the perpetrator had a gun. *Id*. at 8.

[205] Carlson et al. (2016), *supra,* at 8.

[206] *See, e.g.*, Carlson & Carlson, *An Evaluation of Lineup Presentation, Weapon Presence, and a Distinctive Feature Using ROC Analysis*, 3 J. Appl. Res. Mem. & Cog., 45, 46 (2014) ("It is important to use ROC analysis to determine whether the presence of a weapon reduces discriminability, or whether it simply makes eyewitnesses less likely to choose . . . ."). *See* National Academy of Sciences, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 82-87 (2014), for a discussion of ROC (Receiver Operating Characteristics).

There are ethical restrictions on human-subject research that prevent researchers from using violence or the threat of violence. Laboratory studies use videotapes or staged scenarios.[207] While a crime may be portrayed in a videotape, a violent crime cannot be enacted in staged scenarios. How students observe and react under these circumstances are likely different from witnesses to a real crime.[208] "For a laboratory witness," for example, "the presence of a weapon represents no threat and may become a focus of attention due to the orientation of the camera or some other feature that causes the weapon to attract attention."[209] Moreover, participants in an experiment, unlike witnesses to a crime, know that there are no consequences for their identification choices.[210] The external validity of this sort of laboratory research has not been established.

Some weapon focus studies force the subjects to choose, yes or no — sometimes for the sole purpose of increasing identification error.[211] "I don't know" or "I am not sure" or "I didn't get a good enough look at his face" are not options in many laboratory studies even though, in real life, that is how uncertain witnesses often respond. Such responses might, for example, limit the use of an identification procedure in the first place or the use of the witness for identification at trial. In actual cases, police cannot force a witness to choose, and, without more, prosecutors would not rely solely on an uncertain witness for identification at trial.

The presence of a weapon affects feature accuracy much more than identification accuracy.

Many laboratory studies also examine "feature accuracy," the ability to accurately describe the perpetrator.[212] By contrast, "identification accuracy" refers to a witness's ability either to select the perpetrator when he is in the lineup or to reject the lineup when he is not there.[213] "Given its relevance to the criminal justice system, empirical studies need to address why weapon presence appears to be clearly detrimental to feature accuracy (*e.g.*, recall) but only sporadically affects identification accuracy"

---

[207] *See* Pickel et al. (2008), *supra,* at 310 ("Almost all existing studies on weapons focus [are limited to using videos of mock crimes], and those that do not [have this limitation] are restricted instead by confounding variables."). Sometimes these tests are completed remotely online.

[208] *See, e.g.*, Carlson & Carlson, *A Distinctiveness-Driven Reversal of the Weapon-Focus Effect*, Appl. Psychol. Crim. Just., 36, 49 (2012) ("[O]ur findings should not be applied directly to real world situations . . . especially because we could not replicate (nor did we want to, due to ethical constraints) the kind of fear and stress inherent to real eyewitness situations involving a weapon.").

[209] Yuille et al. (2010), *supra,* at 243.

[210] In one study, citizens and police officers were shown a video that simulated a crime, with the viewer as the victim of a robbery. Both citizens and police officers identified the correct perpetrator at the same rate. But police were less likely to make false identifications in target present lineups. This led the researcher to conclude, "police might have a greater knowledge of the criminal justice system . . . and the penalties associated with a wrong decision than citizens. This may inhibit them from guessing when they do not think they will be accurate." DeCarlo (2010), *supra,* at 65-66 (citing Hulse & Memon, *Fatal Impact? The Effects of Emotional Arousal and Weapon Presence on Police Officers' Memories for a Simulated Crime*, Legal & Criminol. Psychol., 313 (2006)). There were no consequences to identifying or not identifying a fake suspect, but unlike the police, the citizens may have had less reason to think about real-world consequences absent any attempt by the researchers to create realistic identification incentives. *See* DeCarlo (2010), *supra,* at 89.

[211] *See, e.g.*, Mitchell et al. (1998), *supra,* at 297 (Mitchell conducted two studies and, after the first study yielded no weapon focus effect where participants had a "don't know" option, they ran the second test—this time forcing participants to make an identification.).

[212] *See, e.g.*, Fawcett et al. (2013), *supra,* at 48.

[213] Fawcett et al. (2013), *supra,* at 48.

(*e.g.*, recognition).[214] The ability to describe the perpetrator is, of course, important to the police's ability to locate a suspect whom they can ask the witness to identify.[215] A description alone, however, would not support a prosecution.

When researchers combine feature accuracy with identification accuracy in reporting their results, it dramatically inflates the weapon focus effect.[216] Suggesting that these results reflect identification accuracy would be wholly misleading.

Average accuracy is meaningless in the courtroom.

To the extent that laboratory weapon focus studies ask subjects to assess their confidence at all, they apply an inappropriate statistical measure (*e.g.*, point-biserial correlation coefficient) to this data.[217] It does not matter in the courtroom how accurate eyewitnesses are *overall* when a weapon is used or not. Experience teaches that some witnesses do not identify a suspect at all, some identify the suspect with low level of confidence, and some identify a suspect with a high level of confidence. Putting them all together says nothing about the accuracy of eyewitnesses who expressed a high level of confidence at the time of the initial *suspect* identification and who are likely to testify at trial. Until laboratory researchers re-analyze their own data on weapon presence and absence using the confidence-accuracy characteristic, it is impossible to draw any meaningful conclusions about the accuracy of eyewitnesses who express a high level of confidence although current research suggests that it is high.

Using the point-biserial correlation coefficient to analyze the weapon focus effect on the accuracy of high-confidence eyewitnesses is like using their team's average to assess Michael Jordan or Kareem Abdul-Jabbar's scoring ability. The average says nothing about how good the best are.

In the laboratory, the weapon focus effect on identification accuracy is small.

In *Henderson*,[218] the court cited a 1992 meta-analysis which found that weapon focus reduces overall "identification accuracy by 10%." This reduction has been described by the author as "small" or "not of great magnitude."[219] In fact, of the 19 studies used in that meta-analysis, 13 showed "no significant differences in the weapon-present and -absent conditions."[220] Since the majority of the weapon focus effect literature "includes correct, but not false identification rate,"[221] it is possible that this "small" effect is caused by eyewitnesses who focused on the weapon and, therefore, did not choose anyone. If

[214] Fawcett et al. (2013), *supra,* at 44. *Accord* Pickel (1999), *supra,* at 310-311 ("Clearly, the presence of a weapon has a smaller effect on witnesses' ability to identify the perpetrator than on their ability to describe him or her accurately"); Steblay (1992), *supra*, at 413 ("Overall the size of the effect was small (.13) for the dependent measure of lineup identification and moderate (.55) for feature accuracy").

[215] *See* Pickel (1999), *supra,* at 310-11 ("[T]he description [of a perpetrator] may be crucial; in cases which the identity of the perpetrator is unknown, the investigation might begin with the description given by the witness, and it might not progress far if that description is incomplete or incorrect.").

[216] *See*, *e.g.*, Fawcett et al. (2013), *supra,* at 47 (discussing confounding identification and feature accuracy).

[217] *See* Confidence and Accuracy.

[218] *State v. Henderson, supra*, 27 A.3d at 905, *citing* Steblay (1992), *supra*, at 420.

[219] Steblay (1992), *supra,* at 417, *id*. at 420.

[220] Steblay (1992), *supra,* at 414. *See* Pickel (1998), *supra,* at 288 ("[W]eapon focus was obtained only with witnesses' descriptions of the target, and not with line-up identification."). *See* text at note 5 *supra*.

[221] Carlson & Carlson (2014), *supra,* at 4.

this is true, an instruction based on statistics that included people who do not choose would mislead the jury in weighing the accuracy of an eyewitness who did choose.

Even in the Carlson et al. (2016) study, where there was an increase in false identifications, it was very small compared to the large decrease in the correct identifications.[222] Moreover, subjects who were highly confident were correct nearly 100% of the time whether a weapon was present or absent. Based on these results, the authors reached the following conclusion: "Should police be skeptical of an eyewitness's identification if they know that the crime involved a weapon? Or should they trust the identification, especially if the eyewitness assigns a high degree of confidence immediately afterward? Our calibration analyses revealed that, not only did the visible weapon not harm the CA relationship, it actually improved it. . . . Based on the CA results from our experiment, we tentatively argue that police could potentially trust an eyewitness who chooses from a lineup and then immediately supports this decision with high confidence."[223]

There is conflicting data in laboratory studies about how the length of exposure affects the weapon focus effect.

The 1992 Steblay meta-analysis of laboratory studies concludes that the weapon focus effect reliably occurs "particularly in crimes of short duration in which a threatening weapon is visible."[224] The analysis does not say how short "short" is. And the *Henderson* instruction does not either. It would be impossible for a jury to apply the instruction without this information even if it applied to false identifications. The 2013 meta-analysis by Fawcett et al. found that the weapon focus effect was diminished if a witness was exposed to the weapon for "short" durations, ten seconds or less, and for "long" durations, more than 60 seconds.[225] However, the meta-analysis also found that there was a "surprising . . . lack of any systematic exploration of exposure duration within the experimental weapon focus literature."[226] Given the absence of adequate research on how the length of exposure affects weapons focus, recommendations based on the ambiguous terms "short" and "long" should not be made.

---

[222] Carlson et al. (2016), *supra*, at 5 (estimated false suspect ID rate 12% when a weapon was absent and 14% when a weapon was present; correct suspect ID rate averaged 50% when a weapon was absent and 25% when a weapon was present).

[223] Carlson et al. (2016), *supra,* at 8. *See* Carlson & Carlson (2014), *supra,* at 51 ("It is possible that the largely absent influence of false IDs could be due to a floor effect, as false ID rate was low across conditions."). Of course, if there were other evidence in the cases, even a low or moderate level of confidence would have some probative value.

[224] Steblay (1992), *supra,* at 421. "[T]he longer the duration, the more time the witness has to adapt to the presence of a weapon and focus on other details." *State v. Henderson*, *supra*, 1 A.3d at 905 (citing testimony of Dr. Gary Wells).

[225] Fawcett et al. (2013), *supra,* at 49; *id*. at 51 ("It could be that a weapon requires a certain period of exposure to ensure that the weapon is perceived and captures attention. With prolonged exposure, the shock of the weapon could fade and the viewer could then begin directing their attention away from the weapon to get a better analysis of their surroundings. This relationship remains a mystery at this time . . ."). The authors apparently choose these amounts of time arbitrarily.

[226] Fawcett et al. (2013), *supra*, at 57. With respect to the duration of the event generally, "[t]here is no measure to determine exactly how long a view is needed to be able to make a reliable identification. Dr. Malpass testified that very brief but good views can produce accurate identifications, and Dr. Wells suggested that the quality of a witness' memory may have as much to do with the absence of other distractions as with duration." *State v. Henderson*, 27 A.3d at 905.

Moreover, the research that purports to find a weapon focus effect often equates the presence of the perpetrator with the presence of the weapon.[227] One study, however, found no weapon focus effect on memory for information when the perpetrator was seen before he displayed a gun.[228] In other words, if the eyewitness saw the perpetrator before the weapon was displayed, the weapon focus effect did not decrease the eyewitness's ability to identify the perpetrator.

To the extent that other studies confound presence of the perpetrator with presence of a weapon, they are of limited ecological validity as well.

The weapon focus effect may be limited by an eyewitness's proximity to a perpetrator.

In an effort to explain why various field studies have not found a significant weapon focus effect, researchers have theorized that "close" proximity to the armed perpetrator may reduce the weapon focus effect.[229] This is particularly true for crimes such as robberies or sexual assaults where the assailant must be close to the victim to effectuate the crime.[230] Yet the laboratory studies do not address whether the weapon focus effect is offense sensitive and whether proximity may be a factor in is applicability.

The weapon focus effect may depend on the individual qualities of the eyewitnesses.

Research has not taken into account that, to the extent that it exists, the weapon focus effect may be entirely dependent on individual experiences, characteristics and perceptions.[231] The 2013 meta-analysis concluded that "[l]ittle work has been done investigating how individual differences shape eyewitness reactions to the [weapon focus effect]."[232]

There is no scientific consensus regarding the cause of the weapon focus effect. Researchers focus on two theories: "arousal theory" and "novelty theory." They posit that eyewitnesses' fear (the "arousal theory") or surprise (the "novelty theory") causes them to focus on the weapon and, therefore, decreases their ability to encode and remember information other than the weapon.[233] The problem is that whether a person is afraid or surprised or neither depends upon the particular eyewitness's individual characteristics and his exposure to and familiarity with weapons in the relevant context. Most laboratory studies test a fairly homogenous population – undergraduates.[234] Even in this context,

[227] *See* Fawcett et al. (2013), *supra*, at 57.
[228] *See* Mitchell et al. (1998), *supra*, at 295 ("This finding supports the contention that the obtained [reduction] is in fact an encoding phenomenon occurring as a function of exposure to, and affecting only details seen at the same time as, the experimental item.").
[229] *See, e.g.*, Mitchell et al. (1998), *supra*, at 44 (citations omitted).
[230] *See* Cooper et al. (2002), *supra,* at 181.
[231] *See, e.g.*, Fawcett et al. (2013), *supra*, at 59. Indeed, this meta-analysis excluded one article from the meta-analysis "on the basis that it used a special population familiar with weapons and responding to events of a violent criminal nature." *Id.* at 46.
[232] Fawcett et al. (2013), *supra*, at 59. *See* Yuille et al. (2010), *supra,* at 243 ("Although more research is needed before any broad conclusions can be drawn, the preliminary results from research with eyewitnesses suggest that laboratory witness finding on the weapon focus effect are not generalizable to actual witnesses.").
[233] *See* Mitchell et al. (1998), *supra*, at 289-292.
[234] *See* Mitchell et al. (1998), *supra,* at 300

"laboratory studies using slides and videotapes sometimes do not produce a reliable weapon focus effect."[235] One researcher explained this inconsistency in the laboratory as follows:

> After all, with all the exposure to weapons provided by the media today it is likely that most college-aged students are becoming desensitized to exposure of weapons in a media-like format. It may be tempting then to leap to the general assumption that a weapon should be expected to be more robust in a real-world eyewitnessing situation. But, it is not clear that such an assumption is necessarily valid.[236]

Nonetheless, the weapon focus effect is insignificant in the laboratory when students expect a weapon to be present -- such as at a gun range or on a police officer.[237] Moreover, one of the few studies involving police officers, who presumably have more experience with weapons and crime,[238] the presence of a weapon actually had a *positive* effect.[239] This instruction, however, would apply equally to all eyewitnesses, regardless of their experience, background, or familiarity with the context of the crime.

Laboratory research on the weapon focus effect is evolving.

Current laboratory research has not reached the level of real-world application by its own standards. Beyond the recognized limitations of laboratory circumstances, social scientists are now calling into question how social scientists have measured weapon focus results. More specifically, a 2014 study by Carlson & Carlson found that most weapon focus effect studies "failed to measure false identification rates" and used ambiguous "diagnosticity" ratios, which merely divide correct identifications by incorrect identifications (including non-identifications).[240] In other words, previous studies, including meta-analyses, combined identifications, regardless of confidence level, even though, "[u]nder ordinary circumstances, these low confidence [identifications] would not play a significant role in a courtroom (i.e., they would be excluded from consideration before reaching that point)."[241]

A few studies are beginning to use more helpful measurements, such as the Receiver Operating Characteristic analysis ("ROC analysis"). ROC analysis measures correct identifications and false identifications based on confidence levels.[242] The limited number of studies that have used ROC analyses are beginning to show very different results than previous studies. More specifically, unlike earlier meta-

[235] *See* Mitchell et al. (1998), *supra,* at 300.
[236] Mitchell et al. (1998), *supra,* at 300; *see also id.* ("what is novel in one case may not be in another (and what is novel to one person may not be to another).").
[237] *See,* Pickel (1999), *supra*, at 309.
[238] Mitchell et al. (1998), *supra*, at 300 ("Might some individuals, namely those regularly exposed to weapons for extended periods of time (*e.g.* police, gang, members, gun dealers), find a weapon less salient/novel/surprising than others?"); *See also id.* at 299 ("[A]n effect completely analogous to the oft-cited weapon focus effect can be reliably produced in a laboratory without a weapon, while inclusion of a weapon does not guarantee a weapon focus effect will occur.").
[239] *See* Hulse & Memon (2006), *supra,* at 322 ("more correct at line-up decisions" when guns used).
[240] Carlson & Carlson (2014), *supra,* at 45.
[241] Gronlund, Wixted & Mickes, *Evaluating Eyewitness Identification Procedures Using Receiver Operating Characteristic Analysis*, 23 Current Directions Pyschol. Sci., 3, 5 (2014).
[242] Gronlund et al. (2014), *supra*, at 4.

analyses, the 2014 study by Carlson & Carlson found that weapon focus did not affect correct identification rates at all.[243]

Going one step further, the use of the confidence-accuracy characteristic ("CAC analysis") in assessing the weapon focus effect, now being studied, promises to provide an even more reliable assessment of how confidence tracks accuracy when a weapon is or is not present.

Furthermore, the 2012 study by Carlson & Carlson found additional nuances to the weapon focus effect in the laboratory when measuring identifications by confidence level. The weapon focus had, for example, a positive, neutral, and negative effect depending on the type of weapon (shotgun or beer bottle) or whether the perpetrator had a distinctive feature (albeit a face-sticker).[244] In some instances, the weapon condition *increased* the correlation between confidence and correct identifications.[245] Thus, it appeared that identifications were more reliable when a perpetrator had a distinct facial feature or when a shotgun was involved. As the authors acknowledge, these studies have weak ecological validity[246] and earlier studies are even less ecologically valid. Recommendations based on them would not be scientifically sound.

Conclusion

The weapon focus effect as summarized in the *Henderson* instruction overstates the laboratory results, ignores the limitations proclaimed by the researchers and does not take into account later research. It does not account for the fact that the presence of a weapon does not increase false identifications of "innocent suspects or targets." Despite the name "weapon focus," very few actual weapons have been studied. Moreover, weapon focus effect has been shown to interact with other variables, even positively, but few interactions have actually been tested.[247] "[M]uch of the previous evidence for the [weapon focus effect] was derived from the same mock video," which depicted a crime from the third-person point-of-view.[248] These studies should not be assumed to apply to real victims and witnesses. Some researchers recognize that current knowledge of the weapon focus effect is far from complete.[249] Existing research does not support a recommendation that the presence of a weapon adversely affects the identification of suspects in real cases.

[243] *See* Carlson & Carlson (2014), *supra,* at 45 (while they did find a weapon focus effect on false identification rates, Carlson & Carlson "encourage" further research and "caution" against the ecological validity because it "could be driven by particular qualities of the mock crime video" used).
[244] *See* Carlson & Carlson (2012), *supra*, at 48.
[245] Carlson & Carlson (2012), *supra,* at 47-48.
[246] Carlson & Carlson (2012), *supra,* at 49 ("[O]ur findings should *not* be applied directly to real world situations" (emphasis added)).
[247] *See, e.g.*, Carlson & Carlson (2012), *supra, at 49.*
[248] Carlson & Carlson (2012), *supra*, at 49 (citing four studies).
[249] *See, e.g.*, Fawcett et al. (2013), *supra,* at 62 ("Although our review and meta-analysis offer much in the way of understanding the nature of weapon focus in real-world crime, it is more a starting point than a final synopsis on the subject.").

SUMMARY

CROSS-RACE EFFECT

Accepting that people may have greater difficulty identifying a person of a different race than identifying a person of their own race in the laboratory setting, it does not mean that witnesses are more likely to identify an innocent suspect of a different race than the same race as themselves, especially with high confidence.

Many laboratory researchers have concluded that there is an "own-race bias" or "cross-race effect" that affects real eyewitnesses, making them less accurate in identifying a person of another race than their own. This conclusion, however, is faulty.

- A racial difference between a target and a subject does not increase the risk of a false identification.
- The cross-race effect has not been found in field or archival studies. There are fewer cross- race identifications, but, of those made, no fewer accurate ones.
- There is no difference in reliability for high-confidence same- and cross-race identifications and only a small difference for moderate-confidence identifications.
- Angry, threatening, or powerful faces – like those most likely to be exhibited in violent crimes – eliminate the cross-race effect.
- Residing in an integrated community, contact with people of other races, having an influential cross-race person in one's life, or having a motive to identify a cross-race person may reduce or eliminate the cross-race effect.
- When researchers report that there is a significant difference, they often mean statistically significant. It does not mean that the effects are practically significant. In many studies, differences in same- and cross-race accuracy are too small to have a practical effect.
- Much of the laboratory research on the cross-race effect has been done in facial recognition studies where exposure times are very small (*e.g.*, .5 milliseconds to 5 seconds; 3 seconds to 6 seconds). Such small amounts of time are unlikely to generate an identification in real cases.
- Significant cross-race effects found in the laboratory when subjects view a single static photograph disappear when they view a moving face. Because a perpetrator's face in real life is not static, cross-race effects are not likely to be present.
- It is not clear whether the photographs used in cross-race studies span the continuum of persons who are identified as one race or another; in one study, "ambiguous" faces were made to be Black or Hispanic by changing hairstyles.
- Facial recognition studies, which predominate in the cross-race literature, cannot and do not measure false identification rates and, therefore, provide no guidance on how a difference in races is likely to affect real witnesses. Researchers recognize that more studies involving lineup tasks are necessary to assess the reliability of the cross-race effect, but few have been done.

The face of America (and the proportion of "minorities") has changed over the past 40 years.

CROSS-RACE EFFECT

> *"[T]here appears to be no substantive difference in rejection accuracy for same-race and cross-race faces . . . ."*

Dodson & Dobolyi, *Confidence and Eyewitness Identifications: The Cross–Race Effect, Decision Time and Accuracy,* Appl. Cog. Pyschol., 113, 120 (2016).

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on the cross-race effect as follows: "Research has shown that people may have greater difficulty in accurately identifying members of a different race. You should consider whether the fact that the witness and the defendant are not of the same race may have influenced the accuracy of the witness's identification."[250]

Analysis

The cross-race effect does not necessarily increase the false identification of innocent suspects (as opposed to reducing the correct identification of guilty suspects).

Accepting that some/many people may have greater difficulty identifying a person of a different race than identifying a person of their own race in the laboratory setting, it does not mean that witnesses are more likely to identify an innocent suspect of a different race than themselves, especially with high confidence. If there is no difference in rejection accuracy, of necessity, there is no difference in false identifications whether the witness and the perpetrator are of the same race or a different race. This means that an innocent suspect of a race other than that of the witness is not more likely to be identified than an innocent suspect of the same race.[251]

The cross-race effect is not found in archival or field studies.

In an archival study, which compared identification accuracy at three levels of corroborating evidence (none, moderate, significant), the authors found that "none of the classic eyewitness factors, race, weapon presence or witness type, produced significant or even marginally significant effects when the identifications were made at field showups."[252] Overall, for lineups, there were more same-race (60%) than cross-race (45%) suspect identifications, but where there was substantial corroboration, there were more cross-race (46%) than same-race (37%) suspect identifications.[253]

---

[250] Press Release, New Jersey Courts, *Supreme Court Releases Eyewitness Identification Criteria for Criminal Cases,* (July 19, 2012), http://www.judiciary.state.nj.us/pressrel/2012/pr120719a.htm; *see State v. Henderson*, 27 A.3d 872, 904 (2011).

[251] *See* Wixted & G. Wells, *The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, Psychological Science in the Public Interest, 10 at 32, Figure 4F (2017) (showing no difference in identification accuracy for same- and cross-race faces for those who were 90-100% confidence and only a slight difference for those who were 70-80% confident.

[252] Behrman & Davey, *Eyewitness Identification in Actual Criminal Cases: an Archival Analysis,* Law & Hum. Behav., 475, 487 (2001).

[253] Behrman & Davey (2001), *supra,* at 481.

In one field study, "no effect of different versus same race of suspects and witnesses was found . . . ."[254] In another field study primarily examining distance, the authors "found no evidence for significant effects of [cross-race] on identification performance."[255] While offering some explanations for why this might be so, the authors stated that "[i]f the cross-race effect is robust, one would have expected it to occur under the conditions of our study."[256] In a recently published field study involving non-blind simultaneous arrays in armed robberies cases, where about two-thirds were cross-race, witnesses rejected the array more often (61%-51%) and identified the suspect less often (26%-36%) when the suspect was of another race and picked a filler at the same rate (12%). The viewing quality and corroborating evidence were higher in cases where suspects were identified than for rejections or fillers.[257] "Of the 295 witnesses who were positive of their selection (the highest level of certainty), 277 (93.9 percent) identified the suspect."[258] It appears then that there are fewer cross-race identifications, but not less accurate high confidence ones.

Although the cross-race effect was not specifically at issue, a recently published field study of armed robberies involved mostly cross-race identifications. Using corroborating evidence to confirm accuracy, high confidence identifications were highly accurate.[259]

Witnesses who are highly confident at the time of the initial identification are highly accurate regardless of race.

Many studies of the cross-race effect (like other factors) measured average accuracy. Because researchers did not ask the participants for confidence statements, or report it when they did, it is impossible to assess whether accuracy was different at different levels of confidence and, therefore, how accurate the high-confidence participants were compared to the low-confidence participants. If one condition yields higher overall accuracy than another, it does not follow that high-confidence identifications in the first condition are more accurate than high-confidence identifications in the second.[260] There are many examples in the laboratory research where overall accuracy differs across conditions but high-confidence accuracy does not differ across those same conditions.[261] This is a critical consideration because, in real life, courts are primarily interested in witnesses who identify a suspect with a high level of confidence. Witnesses who do not identify anyone or who identify a filler, or who

[254] Valentine, Pickering & Darling, *Characteristics of Eyewitness Identification that Predict the Outcome of Real Lineups*, 17 Appl. Cog. Psychol., 969 (2003).

[255] Lindsay, Semmler, Weber, Brewer & Lindsay, *How Variations in Distance Affect Eyewitness Reports and Identification Accuracy*, Law & Hum. Behav., 1, 8 (2008).

[256] Lindsay et al. (2008), *supra,* at 8.

[257] W. Wells, Campbell, Li & Stryker, *The characteristics and results of eyewitness identification procedures conducted during robbery investigations in Houston, TX*, Policing: An International Journal, 601, 609, Table III (2016). The viewing quality and corroborating evidence were higher in cases where suspects were identified than for rejections or fillers. *Id.* "Of the 295 witnesses who were positive of their selection (the highest level of certainty), 277 (93.9 percent) identified the suspect.

[258] W. Wells et al. (2016), *supra,* at 608 ("It was rare for witnesses to select a filler and express the highest degree of confidence.").

[259] Wixted, Mickes, Dunn, Clark & W. Wells, *Estimating the reliability of eyewitness identifications from police lineups,* Proceedings of the National Academy of Sciences, 304, 305 (2016). Note that this field study used different data than the W. Wells (2016) field study.

[260] *See* Confidence and Accuracy.

[261] *See*, *e.g.*, Palmer et al. (2013), *supra,* at 67.

identify a suspect with a low level of confidence are unlikely to be identification witnesses at trial.[262] If real eyewitnesses appropriately moderate their level of confidence by taking into consideration factors like race that might affect their ability to identify a stranger, then the cross-race effect would already be accounted for. It would be misleading, therefore, to suggest that a witness should be given less credence because the perpetrator was of a different race. Indeed, one of the most recent studies of the cross-race effect concluded that "when confidence is collected immediately after the initial identification, and performance is above chance, a high-confidence cross-race identification can be as trustworthy as a high-confidence same-race identification. At the very least, under these conditions, the magnitude of the [cross-race effect] . . . is attenuated when confidence is taken into account."[263]

There is some evidence that in real life people may be less willing or able to make a cross-racial identification than a same-race identification and thus, do not make any identification.[264] This suggests that some people may take into account (consciously or subconsciously) that they may not be as good at recognizing people of another race and act accordingly. There is no reason to assume that witnesses would not moderate their confidence statements when making an identification as well.[265]

One study that asked for confidence statements found that "White participants were much more accurate with White faces than with Black faces. Black participants were also more accurate with White faces than with Black faces."[266] They found that "the relationship between confidence and accuracy was stronger for own-race faces and when the participant makes a positive identification (*i.e.*, says old) [regardless of race]. This is consistent with both laboratory and field studies where own-race confidence is a good predictor of accuracy, but the relationship is much smaller for cross-race identifications. However, the effects are relatively small."[267]

---

[262] Of course, if there is other evidence of identity, a low-confidence eyewitness may testify at trial. If the witness initially says he is, for example 60% confident, the jury can assess that for what it is worth in light of the other evidence in the case. Whether his low level of confidence was attributable to a cross-race effect or something else, would not matter.

[263] Nguyen, Pezdek & Wixted, *Evidence for a confidence-accuracy relationship in memory for same- and cross-race faces*, Q. J. Exp. Psychol., 1, 15 (2016).

[264] Malpass, Zimmerman, Ross & Topp, *Results of the Illinois Lineup Project*, 4 (2006) ("Witnesses shown [blind] sequential cross-race lineups made substantially more non-identifications (64.4%) than those shown [blind] sequential same-race lineups (37.3%) . . . ."); *accord*, Behrman & Davey (2001), *supra,* at 487 ("the tendency to choose the cross-racial suspect from the photographic lineups is lower than the tendency to pick the intra-race suspect in all evidence categories, except the SPV [substantial probative value of corroborating evidence] category").

[265] Sixty-six percent of the respondents in a survey of jury-eligible citizens said that the following statement is *false*: "Generally, eyewitnesses are equally accurate when identifying someone of a different race as when they are identifying someone of their own race." Schmechel, O'Toole, Easterly & Loftus, *Beyond the Ken? Testing Jurors Understanding of Eyewitness Reliability Evidence*, Jurimetrics, 177, 211 (Winter 2006). This is somewhat different from responses to a separate question where 48% of the participants thought that same-race eyewitness were equally reliable to cross-race witnesses, and 36% thought that same-race eyewitnesses were more reliable. *Id.* at 208.

[266] Wright, Boyd & Tredoux, *Inter-racial Contact and the Own-race Bias for Face Recognition in South Africa and England*, Appl. Cog. Psychol., 365, 368 (2003).

[267] Wright et al. (2003), *supra,* at 371. Because the authors did not use a confidence-accuracy characteristic analysis, it is difficult to assess how significant these results are. *See* Glossary and Confidence and Accuracy, which explain calibration. *See also* Jackiw, Arburthnott, Pfeifer, Marcon & Meissner, *Examining the Cross-Race Effect in Lineup Identification Using Caucasian and First Nations Samples,* Canadian J. Behavioural Science, 52, 56 (2008

One other study that included confidence ratings, while asserting that "individuals experience a greater proportion of false recollections for other-race faces – namely incorrect identifications that are made with high confidence," also recognized that:

> As far as the reliability of the CRE, this has only been assessed in two published studies to our knowledge . . . On both occasions, the effect showed only a small degree of test-retest reliability. This low level of reliability, coupled with the absence of relationships to individual measures of memory or potentially relevant attitudes (except for degree of interracial contact . . .) indicates that there is no measure at present that would be forensically useful in predicting which individuals are most likely to manifest a strong CRE in face identification. It remains to be seen whether other potentially useful individual-difference variables can be identified.

Meissner, Brigham & Butz, *Memory for Own- and Other-race Faces: A Dual Process Approach*, Appl. Cognit. Psychol., 545, 563 (2005) (citations omitted).[268]

Angry, threatening, or powerful faces eliminate the cross-race effect.

Most of the cross-race studies present faces with neutral or smiling expressions. However, there may be a primordial instinct that makes cross-race faces "who seem subjectively important (*e.g.*, threatening or powerful) . . . as well-recognized as own-race faces."[269] In order to test this, the authors conducted a recognition study using neutral and angry faces:

> We expected the [own-race bias] to emerge for neutral-expression faces. However, we predicted that expressions of anger would prompt equally accurate recognition of both own-race and cross-race faces, replicating the finding that angry targets eliminate the own-race bias. Critically, because anger induces a highly potent motive for cross-race face individuation, we predicted this would eliminate the own-race bias via an increase in cross-race recognition for both relatively high- and low-experienced participants. In short, we predicted that anger would eliminate the own-race bias, but . . . this motivational effect would not interact with perceiver experience.[270]

This in fact occurred. "[C]ross-race faces displaying expressions of anger, a biologically prepotent facial expression that motivates attentional scrutiny and accurate memory[,] eliminate the own-race bias."[271] Moreover, this was true whether or not the participants had significant contact with people of the other race.[272]

("Although the CRE is known to be a robust effect, the effect sizes reported here are rather small; thus replication of these results is necessary.").

[268] *See also* discussion at note 4 *supra.*

[269] Young & Hugenberg, *Individuation Motivation and Face Experience Can Operate Jointly to Produce the Own-Race Bias*, Social Psychological and Personality Science, 80, 80 (2012) (citations omitted).

[270] Young & Hugenberg (2012), *supra* at 83.

[271] Young & Hugenberg (2012), *supra,* at 84 (citations omitted).

[272] Young & Hugenberg (2012), *supra,* at 84; *see also* Shriver & Hugenberg, *Power, individuation, and the cross-race recognition deficit.* J. Exp. Soc. Psychol., 46, 767 (2010); Ackerman, Shapiro, Neuberg, Kenrick, Becker, Griskevicius, Maner & Schaller, M., *They all look the same to me (unless they're angry): from out-group homogeneity to out-group heterogeneity.* Psychol. Sci., 17, 836 (2006) ("Recognition accuracy for neutral faces showed the out-group

This study and others suggest that laboratory studies of the cross-race effect may have little bearing in the real world where the eyewitness identification of a stranger is most likely to occur in murder, rape, robbery, carjacking, and assault cases. These are all crimes where the assailant is likely to be perceived as angry, threatening or powerful, thereby eliminating the cross-race effect.

Cross-racial contact may reduce or eliminate the cross-race effect.

"[I]ndividuals residing in integrated populations show less of a CRE when compared with same–race individuals residing in more homogeneous populations."[273] "The amount of contact individuals have with other races is also said to be related to recognition of other-race faces, but the evidence is sporadic. . . . A number of studies have shown general support for the differential experience hypotheses, which suggests that the ability to recognize faces of another race is a function not of the absolute amount of contact one has had with members of that race, but the quality of the contact. Face recognition skills may develop from a need to individuate members of populations who are important to be able to identify, such as parents, bosses, and other influential social contacts. . . . The particular features that differentiate individuals within any group will be learned to the extent that it is important to differentiate between individuals in the category in the course of everyday life."[274]

In one study, Korean adults who had been adopted as children by white families were better at recognizing white faces than Korean faces.[275] In another study, "as experience (with other races) increased the magnitude of the ORB decreased, but only when the motive to individuate cross-race faces was activated."[276] In an unusual example, "White basketball fans showed no cross-race recognition effect in recognizing Black faces."[277]

---

homogeneity bias, but this bias was entirely eliminated for angry Black faces. Indeed . . . recognition accuracy was greater for angry Black faces than angry White faces.").

[273] Meissner et al. (2005), *supra*, at 546 (citing studies).

[274] MacLin & Malpass, *Racial Categorization of Faces, The Ambiguous Race Face Effect,* Psych, Public Policy & Law, 98, 99, 100 (2001) (citations omitted). *See* Young, Hugenberg, Bernstein & Sacco, *Perception and Motivation in Face Recognition: A Critical Review of Theories of the Cross Race Effect*, Personality and Social Psychology Review, 116 (2012) ("mere intergroup contact alone appears insufficient to improve CR memory, but higher quality contact may be successful in doing so"); Shriver & Hugenberg (2010), *supra,* at 772 ("Our results show that both positive and negative CR targets can be highly memorable. As minorities achieve increasing social power and economic success, and are seen in increasingly high-status roles, this may attenuate the CRE."); Walker & Hewstone, *A Perceptual Discrimination Investigation of the Own-Race Effect and Intergroup Experience*, Appl. Cog. Psychol., 461, 470 (2006) ("As no own-race bias effect for South Asian participants was observed, no subsequent regression analyses of social contact or individuating experiences was conducted . . . . White participants with higher levels of individuating experience with South Asians showed increased discrimination accuracy for South Asian faces.").

[275] Brigham, Bennett, Meissner & Mitchel, *The Influence of Race on Eyewitness Memory*, In Lindsay, Ross, Read & Toglia, (Eds), HANDBOOK OF EYEWITNESS PSYCHOLOGY: MEMORY FOR PEOPLE, 257, 260 (2007) (citation omitted).

[276] Young & Hugenberg (2012), *supra*, at 82 (Motive to individuate was created by telling the subjects to "pay close attention to what differentiates one particular face from another face of the same-race, especially when that face is not of the same race as you.").

[277] Li, Dunning & Malpass, *Cross-racial identification among European-Americans: Basketball fandom and the contact hypothesis*, unpublished manuscript (1998), *cited in* Da Silva, ACCURACY AND DECISION MAKING CRITERIA IN CROSS-RACE EYEWITNESS IDENTIFICATION: A MORE COMPLEX THAN EXPECTED PHENOMENON (City University of NY. 2008).

The differences between same- and cross-race identification accuracy may not be practically significant.

Although a cross-race effect may be statistically significant, the actual differences may be so small that they do not provide a firm foundation for recommendations that suggest that racial differences may adversely affect the reliability of an eyewitness's identification.[278] For example, in a recent (2016) study on the cross-race effect, researchers found that for witnesses who are 100% confident, the difference in accuracy between same- race and cross-race identifications is 3% (80% accuracy for same-race; 77% for cross-race); in the aggregate (all levels of confidence combined), the difference in accuracy between same- and cross-race identifications also is 3% for whites (48% same-race, 45% cross-race) while for blacks it is larger (60% same-race, 43% cross-race).[279] However, a re-analysis of the data using the confidence-accuracy characteristic, a more informative metric, shows no difference in accuracy for same- and cross- race identifications when witnesses are 90-100% confident (95-96%) and a very small difference in accuracy when witnesses are 70-80% confident (83-86%).[280]

In another study that compared recognition accuracy between sober and moderately intoxicated participants, "[p]aired- samples t-tests showed that in the placebo condition, recognition accuracy was significantly higher for same-race faces (M=.86) compared to different race faces (M =.81). In the alcohol condition, however, there was only a small tendency toward better recognition performance for same-race faces (M = .81) compared to different-race faces (M = .79), and this difference was not reliable."[281] In a third study, "there was a significant main effect of race on accuracy such that participants were more accurate for own-race faces (M = .87) than other-race faces (M = .84).[282] In a fourth study, the authors wrote that "[a]lthough the cross-race effect is known to be a robust effect, the effect sizes reported here are rather small; thus replication of these results is necessary."[283] A "small" – 2% or 3% or 5% -- difference may be statistically significant, but it is not practically significant. Other studies appear to show a larger difference,[284] but given the variability, it is hard to assess the true magnitude the cross-race effect – and whether or how it applies to specific individuals.

---

[278] "When you have a large sample size, very small differences will be detected as [statistically] significant. This means that you are very sure that the difference is real (i.e., it didn't happen by fluke). It doesn't mean that the difference is large or important." www.statpac.com/surveys/statistical-significance.htm. "In normal English, 'significant' means important, while in statistics 'significant' means probably true (not due to chance). A research finding may be true without being important." www.surveysystem.com/signif.htm. *See* Wells, *Scientific Status*, in Faigman, Kaye, Saks & Sanders, MODERN SCIENTIFIC EVIDENCE, 475, 496 (2005) ("Something can be statistically significant and yet have a small effect size.").

[279] Dodson & Dobolyi (2016), *supra,* at 118.

[280] Wixted & G. Wells, *The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, Psychological Science in the Public Interest, 10, 32, Figure 4F (2017).

[281] Hilliar & Kemp, *Now Everyone Looks the Same: Alcohol Intoxication Reduces the Own-Race Bias in Face Recognition*, Law & Hum. Behav., 367, 372 (2010) (other statistics omitted) (Asian and European).

[282] Marcon, Meissner, Fruch, Susa & MacLin, *Perceptual identification and the cross-race effect*, Visual Cognition, 767, 771 (2010) (Hispanic and African American).

[283] Jackiw et al. (2008), *supra,* at 56 (First Nation and White).

[284] One 1997 article said that "[t]he effect is typically a 10% to 15% difference in accurate same-race versus cross-race identifications." Teitelbaum & Geiselman, *Observer Mood and Cross-Racial Recognition of Faces*, J. Cross-cultural Psychol., 93 (1997) (subjects viewed 20 faces for 5 seconds each). A 2001 meta-analysis posited that "the ORB in discrimination accuracy accounted for 15% of the variability across studies." Meissner & Brigham, *Thirty Years of Investigating the Own-Race Bias in Memory for Faces: A Meta-Analytic Review*, Psychology, Public Policy &

Even under the short exposure time/multiple identification tasks in the laboratory, the average accuracy rate is relatively high.[285] With longer exposure times and fewer faces, overall accuracy is likely to rise.

The general acceptance of the cross-race effect was supported by a 2001 study, in which 90% of the 64 experts surveyed were said to agree that "eyewitnesses are more accurate when identifying members of their own race than members of other races."[286] However, "this does not mean that the bias always occurs, that it is large, or that it is equally strong for everyone."[287] Moreover, one of the leading experts in this field observed that "[w]e know that the other-race effect on face recognition exists, but we do not know what basis might exist for predicting that a specific person might be subject to it or exempt from it."[288]

As the exposure time increases, the cross-race effect disappears; the difference in exposure times may be tiny (*e.g.*, 0.5 to 1.5 seconds).

It is generally recognized that "[a]s viewing conditions become more 'optimal,' recognition performance (the ability to correctly identify previously seen faces) improves. Some factors that affect optimal viewing conditions are the amount of time a person has to view the face and the amount of time that lapses between the initial viewing and the test of memory."[289] In the MacLin study, the viewing times were 0.5 seconds and 5 seconds. In the Meissner & Brigham meta-analysis, the mean viewing time was 3 seconds. In a more recent study, the viewing times were even smaller. In the Marcon study, the "CRE increased as the amount of time participants encoded a target face decreased. . . . [S]ignificant CREs were observed at the 100 ms and 500 ms encoding conditions, but were not observed when encoding time was 1000 ms and 1500 ms (statistical formulas omitted)."[290] A "ms" or millisecond is 1/1000 of a second. So the cross-race effect was no longer seen at 1 and 1.5 seconds.

Law 3, 16 (2001). This means that 85% of the variability across studies was due to factors having nothing to do with the effect of race on discrimination accuracy.

[285] *See* Hilliar & Kemp (2010), *supra*, at 376 ("It was the case that recognition accuracy across the experimental conditions was never lower than 75%.") and Table 1 (showing, at most, a three-point difference and as little as zero difference between same- and cross-race accuracy: 79-76, 80-80, 76-73 and 75-74); Marcon et al. (2010), *supra,* at 771 (across all trials, mean accuracy was 85.80%).

[286] Kassin, Hosch & Memon, *On the "General Acceptance" of Eyewitness Testimony Research, A New Survey of the Experts*, Am. Psychol., 405, 410, 411 (May 2001). Forty percent (25/63) thought the statement was very reliable, thirty percent (19/63) thought it was generally reliable, and twenty-five percent (16/63) thought it tended to be reliable. *Id*. at 411, Table 2. The article's claim that over 90% of the respondents judged cross-race bias, among other factors, as "reliable" appears somewhat of an overstatement given that 25% found that the data only tended to favor such an assessment, *id*. at 411, Table 3, and only 72% would testify on this subject. *Id.* at 412, Table 4. Others have questioned the generalizability of the Kassin survey. *See* text at footnotes 129-131, *supra*.

[287] Wright, Boyd & Tredoux, *Inter-racial Contact and the Own-race Bias for Face Recognition in South Africa and England*, Appl. Cog. Psychol., 365, 366 (2003).

[288] MacLin & Malpass (2001), *supra,* at 114.

[289] MacLin, MacLin & Malpass, *Race, Arousal, Attention, Exposure, and Delay: An Examination of Factors Moderating Face Recognition*, 7 Psychol. Pub. Pol'y, 134, 135-136 (2001); *accord* Meissner & Brigham (2001), *supra,* at 19 ("Results indicated that the amount of study time influenced estimates of the own race bias on measures of discrimination accuracy. The direction of the effect indicated that reducing the amount of study time for each face significantly increased the magnitude of the ORB, largely as a result of an increase in the proportion of false alarm responses to other race faces.").

[290] Marcon et al. (2010), *supra,* at 771-772.

These exposure times are tiny, less than those in crimes where an eyewitness is usually able to identify a stranger. If the CRE is significantly reduced or disappears after one, three or five seconds, then what is the basis for an instruction when the exposure time in a real crime is greater, especially for witnesses who are certain of their identification?

The cross-race effect found with exposure to a static face disappears with exposure to a moving face.

As discussed above, most of the studies of the cross-race effect have used a facial recognition, not eyewitness identification, paradigm. This raises the question whether the cross-race effect is found when a face is viewed in motion, as it would be in real life, and not in the single static pose used in facial recognition studies. The question was recently answered in the negative.[291] Researchers compared a single static pose (for 8 seconds), a series of four static poses (for 2 seconds each), and a rapidly moving face (composed of 18 still photos rotating back and forth for 112 ms each). In all of the conditions, 12 Caucasian and 12 Asian faces were displayed to Caucasians and Asians subjects for a total of eight seconds each. Recognition was then tested by displaying 24 faces of each race, half old, half new. The "significant" cross-race effect present in the single static pose disappeared in both of the other conditions.[292] The authors also found that "difference patterns of [CRE] across experiments cannot be attributed to different levels of contact with other-race faces."[293]

It is not clear whether the selected for study span the continuum for each race.

The studies do not describe how they selected same- and cross-race faces for inclusion. "This raises the question of whether the faces used in these experiments to represent people of different races actually represent the full range of facial feature variations that actually exist."[294] "It is a fact of life in many urban environments that persons belong to multiple ethnic, identity, and appearance groups. A person could be described as Hispanic, Black, Mexican, Indian or Asian and have some facial features stereotypically characteristic of other groups."[295] A Caucasian from Sweden may bear little resemblance to a Caucasian from Turkey. In one study, the same "ambiguous" faces were categorized as Hispanic or Black simply by changing "stereotypical" hairstyles.[296] These variations make it inappropriate to apply these studies to real life where a range of faces belong to any given racial or ethnic group or a mixture of them.

---

[291] Zhao, Hayward & Bulthoff, *Face format at encoding affects the other-race effect in face memory*, J. Vision, 1 (2014). These authors used the term "other-race effect (ORE)," but for the sake of consistency and to eliminate confusion with "own-race bias (ORB)," we use the term "cross-race effect (CRE)."

[292] Zhao et al. (2014), *supra,* at 5, 6 & 7.

[293] Zhao et al. (2014), *supra,* at 8. There appears to be a difference in rigid motion and elastic motion that requires further study.

[294] Smith & Stinson, *Does Race Matter? Exploring the Cross-Race Effect in Eyewitness Identification*, in Parks, Jones & Cardi, INTERSECTIONS OF PSYCHOLOGY, RACE, AND LAW, 102, 106 (2008) ("[T]he magnitude of the cross-race effect is probably affected by variations across faces of people of different races. This raises the question of whether the faces used in these experiments to represent people of different races actually represent the full range of facial feature variations that actually exist.").

[295] MacLin & Malpass (2001), *supra*, at 113. The authors said this to caution the police against uncritically accepting "the witness's racial categorization of a perpetrator" because it might "lead to lineups that miss important feature pools on which the choice of lineup fillers should draw." *Id.* Because suspects are usually chosen when there is some corroborating evidence, this danger is not as great as it may appear to laboratory researchers.

[296] MacLin & Malpass*, Racial Categorization of Faces: The Ambiguous Race Effect*, 7 Psychol, Publ. Pol'y & L., 98 (2001).

Facial recognition studies do not measure false identification rates.

A facial recognition study is a kind of list-memory study. In a list-memory study, when the correct rate goes down, the false rate will always go up. This phenomenon is well known and is called the mirror effect. It is an artifact of the list-memory design. It may be useful for investigating the effect of cross-race identifications on discriminability overall, but it does not provide guidance on false identification (or false alarm) rates for laboratory subjects who participate in a more forensically relevant study or real-life witnesses who view a lineup, show-up, or photo array.

The 2001 meta-analysis, based overwhelmingly on facial recognition studies, provides little useful guidance on the implications of racial difference on the accuracy of real life witnesses.

Meissner & Brigham authored one of the most frequently cited meta-analyses on the cross-race effect.[297] In assessing meta-analyses in the field generally (but citing this one for ORB specifically), the National Academy of Sciences found that "none of the reviews met all current standards for conducting and reporting systematic reviews, and few met even a majority of these standards, making assessment of the credibility of their findings problematic. After examining the reviews, the committee concluded that the findings may be subject to unintended biases and that the conclusions are less credible than was hoped."[298]

With this caveat in mind, and acknowledging that this meta-analysis found a "significant [own race bias]," it is worth noting that:

- 91% or 35½ of the 39 articles in this meta-analysis were facial recognition studies. Because facial recognition studies are not informative on the key issue in a criminal trial, that is, how reliable suspect identifications are, this work is of little value. The authors themselves recognized "only a small proportion (3.5 articles or 9%) of the samples involved the use of an identification task, [therefore,] future studies utilizing the lineup paradigm would be valuable." *Id*. at 14. To the best of our knowledge, this has not been done in the intervening 15 years.
- The remaining 9% (or 3 ½) studies all used "a (simultaneous and target-present) lineup identification task." Thus, they provide no information on whether a difference in race affects the false identification rate. The authors recognized that "[i]dentification paradigms are generally more applicable to the eyewitness situation." *Id.*
- Even if the underlying studies were more forensically relevant, the exposure time in all of the studies ranged from 0.12 seconds to 4 minutes with a median 3 seconds.[299] *Id.* at 14. An exposure time of three seconds, more or less, is unlikely to generate the identification of a stranger in most real life cases.
- "The size of the [own-race bias] has significantly decreased over time for measures of discrimination accuracy and proportion of hits, [but] significantly increased over time for

[297] Meissner & Brigham, *Thirty Years of Investigating the Own-Race Bias in Memory for Faces; A Meta-Analytic Review*, Psychology, Public Policy & Law, 3 (2001).
[298] National Academy of Sciences, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 76 (2014).
[299] The authors do not disclose how many studies exceeded the median by only a few seconds. Most of the facial recognition studies we have seen have an exposure time of note more than 5 seconds.

the proportion of false alarms."[300] *Id*. at 20. This is consistent with the mirror effect for list-memory studies and does not indicate that false identifications in real cases are rising.

- "Results indicated that the amount of study time influenced estimates of the [own-race bias] on measures of discrimination accuracy. The direction of the effect indicated that reducing the amount of study time for each face significantly increased the magnitude of the [own-race bias], largely as a result of an increase in the proportion of false alarm responses to other race faces." *Id.* at 9.
- "[I]t is increasingly evident that the contact hypothesis plays a vital role in our conception of the [own-race bias]." *Id*. at 20-21 (citations omitted).[301]

These factors – the use of facial recognition rather than eyewitness identification studies to measure ORB/CRE; the very small amount of exposure time, often in milliseconds; the large number of faces seen in the studies compared to one or two in real life; the reduction in the ORB/CRE over the three decades that the studies spanned, a period of time in which racial contact was increasing and has continued to increase since; and a reduction in false alarms of other race faces with increased study time – are issues that cast doubt on whether and how the identification of a person of another race in real life is affected.

Finally, there is no indication that any of the articles included in this meta-analysis used confidence statements to measure individual accuracy as contrasted with group accuracy. Without such an assessment, which law enforcement is bidden to do in real life, the average accuracy is really meaningless to the issues in a criminal trial.

Conclusion

For the reasons set forth above, recommendations that suggest that an eyewitness might be mistaken because the perpetrator is of a different race would not be based on the full range of research and the extensive variability within it. The research to date does not reflect the circumstances in real cases. In the laboratory studies, the differences are not uniform, they are often tiny, they are moderated by a longer (but still tiny) exposure duration and more extensive contact with other races, and, although the average accuracy rates are very high across the board, they do not distinguish between subjects who are certain and those who are not. There is no evidence that people are more likely to identify an innocent suspect of a different race than of the same race. In studies that evaluate a threatening facial expression – the most likely expression used in a crime of violence – there is no cross-race effect. In the most recent laboratory studies, the cross-race effect was very small. There is no cross-race effect in field and archival studies. And highly confident witnesses are highly accurate irrespective of race. Current research does not support recommendations that fail to incorporate these findings.

[300] 27% of the articles were published or written in the 70s; 33% in the 1980s; and 40% in the 1990s; 15% were not published. NAS at 14.

[301] Although it is generally believed that people who are less prejudiced in their beliefs would be better at cross-race identification, one study suggested the reverse is true. Edlund & Skowronski, *Eyewitness Racial Attitudes and Perpetrator Identification: The Lineup Method Matters*, North American J. Psych., 15, 25 (2008) (When shown a sequential array only, "[h]ighly prejudiced people were more likely to correctly identify the [African American] perpetrator [seen for 1.5 seconds] when he was in the lineup, but were also less likely to falsely identify a foil when the perpetrator was not in the lineup.").

SUMMARY

EXPOSURE TIME, DISTANCE, LIGHTING, RETENTION, DISGUISE, INTOXICATION, AND DESCRIPTION

Researchers have pointed to these factors that, they claim, reduce a witness's ability to accurately identify a perpetrator. The *Henderson* instructions uses phrase like: "less likely to produce an accurate identification", "may not always be accurate," "the higher the risk of misidentification," "tends to be more unreliable," "can reduce the accuracy of an identification," "the greater the possibility a witness's memory will weaken." However, these descriptions are misleading. Research now strongly supports the conclusion that these factors may reduce the number of identifications that are made, but they do not: (a) significantly increase false identifications of innocent suspects, and/or (b) reduce the reliability of identifications that are made, particularly those that are made with high confidence.

Putting everything else aside, ordinary people experience these factors every day. To the extent that research is counter-intuitive, it goes in the opposite direction of what has been presented to the public and the courts up to now.

Exposure time: Research now shows that a people make fewer identifications when the exposure time is short (*e.g.*, 5 seconds), but when they are made with high confidence, they are as reliable as high confidence identifications when the exposure time is long (*e.g.*, 90 seconds).

Distance and Lighting: There is only a handful of studies on distance and/or lighting that reach the unremarkable and intuitive conclusion that identification accuracy decreases as distance increases and lighting decreases. It is not yet clear whether adverse conditions increase false identifications significantly, or at all. However, in one study, in a very dark room (less than 0.3 lux) at a distance of 16.5 feet, 78% of the subjects correctly said that photographs were not those of a coworker (and when the face was illuminated by a flashlight, that percentage rose to 98.8%).

Retention Interval: Ordinary people know from their own experience that memory may fade over time. Current research indicates that the longer the time between the crime and the identification procedure, the less likely it is that a witness will make an identification. Thus, there will be *fewer* identifications. However, those high confidence identifications that are made are *no less accurate* than those made shortly after the crime. This accords with general memory studies that find that traumatic events are more likely to be remembered than every day or neutral events.

Disguise: If a person is wearing a mask that covers his face, it is unlikely that he will be identified later (except if a witness recognizes his eyes, which is more likely to happen if they know each other). Research that characterizes hats and sunglasses as disguises have inconsistent results. In one study, hats and hats and sunglasses did not increase false identifications, whereas sunglasses alone did (by about 1.5%). In another, a baseball cap and dark sunglasses had no significant effect on identification in either the target-present or the target absent lineups. Overall, confidence decreased with disguises.

Intoxication: The handful of studies on intoxication and eyewitness identification do not support the proposition that intoxication reduces accuracy overall, or more importantly, increases false identifications. Two recent studies have found that alcohol decreases hit rates but has no effect on false alarm rates. This means that an innocent suspect is no more likely to be picked by an intoxicated person than a sober one.

Description: A 2014 review of the literature found that despite the clear intuition that witnesses who are better at describing a target should also be better at recognizing it, this relationship has proved to be quite elusive and generally weak. Recommendations based on the assumption that there is a close relationship between description and identification accuracy would be faulty.

EXPOSURE TIME, DISTANCE, LIGHTING, RETENTION, DISGUISE, INTOXICATION AND DESCRIPTION

Introduction

The series of instructions proposed in *Henderson* that have to do with the circumstances under which a witness views the perpetrator are not supported by the research. The current jury instruction (9.210) informs the jury that it may consider:

> The witness's opportunity to observe the criminal acts and the person committing them, including, but not limited to, the length of the encounter, the distance between the various parties, the lighting conditions at the time, and the witness's state of mind at the time of the offense. Any subsequent identification and the circumstances surrounding that identification, including the length of time that elapsed between the crime and the identification . . . .

This instruction clearly outlines the factors that the jury may take into consideration in weighing the accuracy of an identification. It is short and balanced. It allows the jurors to apply their own experiences and common sense to an identification.

Research does not support comments like: "less likely to produce an accurate identification," "may not always be accurate," "the higher the risk of misidentification," "tends to be more unreliable," "can reduce the accuracy of an identification," "the greater the possibility a witness's memory will weaken." All of these statements are at best ambiguous and misleading; at worst – and there is good reason to believe the worst – they are wrong. They provide no yardstick against which a particular witness's testimony can be measured. They provide no meaningful guidance to the courts or the public, scientific or otherwise.

All of this comes down to a single point: exposure time, distance, lighting, intoxication, delay, disguise, and description are all factors that ordinary people are perfectly capable of assessing without the assistance either of an expert witness or biased instructions that may prejudice the jury without giving them appropriate guidance.[302]

EXPOSURE TIME OR EXPOSURE DURATION

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on exposure time as follows: "The amount of time an eyewitness has to observe an event may affect the reliability of an identification. Although there is no minimum time required to make an accurate identification, a brief or fleeting contact is less likely to produce an accurate identification than a more prolonged exposure to the perpetrator. In addition, time estimates given by witnesses may not always be accurate because witnesses tend to think events lasted longer than they actually did."

[302] "It is doubtless true that from personal experience and intuition all jurors know that an eyewitness identification can be mistaken, and also know the more obvious factors that can affect its accuracy, such as lighting, distance, and duration." *People v. Murphy*, No. B238006, 2013 WL 2242449, at *3 (Cal. Ct. App. May 20, 2013), review denied (July 31, 2013).

Analysis

It probably goes without saying that the longer a person has to observe a stranger, the more likely it is that s/he will be able to identify him/her. We intuitively know this to be true. There is, however, a difference between the *ability* to identify a stranger, if one can, and the *reliability* of an identification that is made. The less time a witness has to observe, the less likely it is that the witness will identify anyone. However, the accuracy of those who actually *make* a high-confidence identification is not impaired by a short exposure to a stranger's face.

The most recent study on exposure time found that "while participants in [the short exposure] condition were less likely to make relatively high-confidence IDs, when they did, they were as accurate as the high-confidence IDs from the long exposure condition. . . . [Thus,] a high-confidence ID made from the 5 [second] condition was as likely to be correct as a high-confidence ID made from the 90 [second] condition."[303] Because only highly confident witnesses are likely to testify at trial, there is no basis to assert that their identifications may be less accurate because of a short exposure time. This research undermines the assertion that a high-confidence identification made after a short exposure time is unreliable. Moreover, for a 90-second exposure time, a medium-confidence identification was accurate over 90% of the time, and a low-confidence identification was accurate just less than 90% of the time.[304]

Even if older studies themselves were reliable, the phrase "brief or fleeting" is ambiguous. Does it mean one second or less, ten seconds, a minute or more? Existing research gives no reliable demarcation. In a considerable amount of eyewitness identification research, the subjects are exposed to faces for milliseconds – and still achieve high rates of identification accuracy. For example, in one study "[r]ecognition performance was superior for longer exposure times (5 seconds) than for brief presentation (0.5 seconds) . . . ."[305]

One meta-analysis looked at 33 studies that measured the difference in accuracy between two exposure times, one "short" and one "long."[306]

---

[303] Mickes, *Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory*. J. Appl. Research Mem. & Cog., 93, 96 (2015).

[304] Mickes (2015), *supra*, at 95, Figure 1C. For the five-second exposure time Identification accuracy was less by about 7% for those with medium confidence and 15% for those with low confidence.

[305] MacLin, MacLin & Malpass, *Race Arousal, Attention, Exposure, and Delay: an Examination of Factors Moderating Face Recognition*, 7 Psychol. Pub. Pol'y, 134, 140 (2001). This is a typical facial recognition study in which the subjects viewed 20 Hispanic and 20 Black faces initially for 0.5 or 5 seconds each and, after half of the photographs were replaced, viewed the new set for 5 seconds each to identify whether they were "old" or "new" faces. In another facial recognition study, the cross- race effect was eliminated when duration increased from 100 ms and 500 ms to 1000 ms and 1500 ms. Marcon, Meissner, Frueh, Susa & MacLin, *Perceptual identification and the cross-race effect*, Visual Cognition, 767, 771-772 (2010).

[306] Bornstein, Deffenbacher, Penrod & McGorty, *Effects of exposure time and cognitive operations on facial identification accuracy: a meta-analysis of two variables associated with initial memory strength*, Psychol., Crime & L., 473, 477 (2012). Citing this meta-analysis, among many others, the National Academy of Sciences [NAS] commented that "none of the reviews met all current standards for conducting and reporting systematic reviews, and few met even a majority of these standards, making assessment of the credibility of their findings problematic. After examining the reviews, the committee concluded that the findings may be subject to unintended biases and that the conclusions are less credible than was hoped." National Academy of Sciences, Committee on Scientific

- Thirty-one of the studies were facial recognition studies; two were eyewitness identification studies.[307]
- "[T]he difference in exposure times between long and short durations ranged from 0.7 to 3570 [seconds or one hour]" with a median difference of 4.7 seconds[308] This means that the *difference* in exposure times for half of the studies was less than 4.7 seconds.
- Extrapolating from the data on Table 1, the ranges below the median would have been along the lines of: 1.25 to 3.75 seconds (Wallace); 2 to 4 seconds (Brigham); 1 to 5 seconds (Malpass); 3 to 6 seconds (Meissner); 0.2 to 1.5 seconds (Semmler & Brewer).[309] Some people might interpret the longer exposure times in these studies – 1.5 seconds, 3.75 seconds, 4 seconds, 5 seconds, 6 seconds – as "brief or fleeting" and yet identification accuracy for these exposure times was better than at the shorter exposure times.
- "Twenty-eight of the 33 studies had what [the authors classified] as 'short' shorter durations ranging from a few tenths of a second to 12 seconds, and longer durations of no more than 45 seconds. Twenty-five of these 28 studies showed strong positive effect sizes . . ."[310]
- "Four of the remaining five studies had what we classify as 'long' shorter durations of 45-90 [seconds] and longer durations of 180-480 [seconds]. The effect sizes for these four studies were negligible. . . ."[311]

The authors concluded, "[o]nce the exposure time for a *briefly* presented face exceeds 30 [seconds] or so, any further increase would have to be quite substantial to produce a further improvement in performance."[312] As significantly, they described exposure times of 5 to 30 seconds as "moderate," not brief.[313]

A short exposure time (whatever that may be) may result in fewer witnesses who are able to make an identification, but it appears that the high-confidence identifications among them are no less accurate than in a long exposure time. In a recent study that compared accuracy between 5-second and 90-second exposures to the target, the original authors found that "accuracy clearly increased with confidence in all exposure and retention interval conditions. This was particularly evident in the upper half of the confidence scale, and especially at the upper end of the scale (i.e., 90%-100% vs. 70%-80%)

Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 75-76 and n.9 (2014) (footnote omitted).

[307] Bornstein et al. (2012), *supra*, at 482 ("A moderator analysis comparing EWID studies and face recognition studies was not possible, due to there being only two studies that had been conducted within the EWID paradigm.") *See* Memon, Hope & Bull, *Exposure duration: Effects on eyewitness accuracy and confidence*, British J. Psychology, 339, 340 (2003) ("There is only a handful of studies that have manipulated the exposure duration in an eyewitness context. . . . While researchers tend to report the length of the stimulus event, they rarely report actual exposure to the target's face.").

[308] Bornstein et al. (2012), *supra*, at 477. There was one study at the 4.7 mark. The difference in exposure times for the 16 studies below the median ranged from 0.7 to 4 seconds. The difference in exposure times for the 16 studies above the median ranged from 10 seconds to one hour. *Id.* at 478, Table 1.

[309] Unfortunately, the article does not give the ranges, but only the differences between the "short" and the "long" exposure.

[310] Bornstein et al. (2012), *supra*, at 484.

[311] Bornstein et al. (2012), *supra*, at 484.

[312] Bornstein et al. (2012), *supra*, at 484 (emphasis added).

[313] Bornstein et al. (2012), *supra*, at 485.

confidence). Together, these results suggest . . . confidence remained a useful indicator of accuracy in all experimental conditions."[314] A second author, reanalyzing the data from the first, found that "not surprisingly, memory was better (discriminability was higher) when exposure duration was longer. . . ." Indeed, the data "indicate that the participants appreciated the effect that exposure time would have on their memory and compensated for it by appropriate adjusting their confidence, particularly at the high-confidence end of the scale. . . ."[315] The key point here is that "while participants in [the 5 second exposure] condition were less likely to make relatively high confidence IDs, when they did, they were as accurate as high confidence IDs from the long exposure condition."[316]

It appears to be not true then, that "the amount of time an eyewitness has to observe an event may affect the reliability of an identification" or "a brief or fleeting contact is less likely to produce an accurate identification than a more prolonged exposure to the perpetrator." There may be both fewer identifications and fewer high confidence identifications with a short exposure time, but high confidence identifications are just as reliable as those made with a long exposure time.

One study concludes that people may over-estimate the time of short events and underestimate the time of long events.[317] Putting aside the fact that it does not appear to have been replicated, this study is of little practical significance. First, it appears that differences in time estimates do not affect the amount of information that is correctly recalled.[318] Second, many crimes these days are captured by surveillance cameras, thereby documenting the duration of an event independent of witness estimates. Third, even if a witness estimated the exposure duration incorrectly, it does not matter: high confidence identifications are just as reliable in short as in long exposure durations.

DISTANCE AND LIGHTING

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on distance and lighting as follows: **Distance:** A person is easier to identify when close by. The greater the distance between an eyewitness and a perpetrator, the higher the risk of a mistaken identification. In addition, a witness's estimate of how far he or she was from the perpetrator may not always be accurate because people tend to have difficulty estimating distances. **Lighting:** Inadequate lighting can reduce the reliability of an identification. You should consider the lighting conditions present at the time of the alleged crime in this case.

---

[314] Palmer, Brewer, Weber & Nagesh, *The Confidence-Accuracy Relationship for Eyewitness Identification Decisions: Effects of Exposure Duration, Retention Interval and Divided Attention*, Am. Psych. Assn., 55, 61 (2013).

[315] Mickes (2015), *supra*, at 95, 96.

[316] Mickes (2015), *supra*, at 96.

[317] Yarmey, *Retrospective Duration Estimations for Variant and Invariant Events in Field Situations*, Appl. Cog. Psychol., 45, 53 (2000) ("duration estimations for relatively short events were overestimated whereas longer events tended to be relatively accurately estimated or underestimated. The definitions of 'short' and 'long' events are arbitrary, . . . [but] the cut-off point for variant events appears to be between 13 and 16 minutes. For invariant events the cut-off point fell between 2.5 and 13 minutes. The mean percentage error for shorter variant activities ranged between 25% for an event lasting 13 minutes in duration [*e.g.*, 17 minutes] to 115% overestimations for an event lasting 16 seconds [*e.g.*, 34 seconds]."") (parentheticals omitted).

[318] Loftus, Schooler, Boone & Kline, *Time Went by so Slowly: Overestimation of Even Duration by Males and Females*, Appl. Cog. Psychol., 3, 6 (1987) ("time estimation was unrelated to amount of free recall"); *id*. at 7 ("The correlation between time estimation and accuracy of the remaining test items was -0.02 and not significant").

Analysis

Like exposure time, distance and lighting are matters that are well within the experience of the ordinary person and research does little to further elucidate these factors. Because distance and lighting generally can be verified objectively, whether the witness accurately estimated either is irrelevant.

Distance and lighting appear to have an inverse relationship. That is, as distances decrease and lighting increases, the proportion of accurate identifications increases; conversely as distances increase and lighting decreases, the proportion of accurate identifications decreases. But it is not an absolute:

> [T]he likelihood of correct identifications goes up when the distance decreased from 40 to 7 meters (131 to 23 feet); below 7 meters, little is gained. When illumination increases, recognition is also improved, but the major part of gain is between 0.3 (night with full moon) and 10 lux (urban area with bright street lights). Beyond 10 lux little is gained."[319]

Indeed, at the longest simulated distance of 131 feet and simulated illumination of 3000 lux (cloudy day) in the recognition test (Wagenaar), 29% correctly identified a stranger, and 9% incorrectly identified a filler.[320] At the shortest simulated distance measured of 10 feet with the same simulated illumination, 86% correctly identified a stranger and only 2% incorrectly identified a filler.[321]

Another study (Lindsay) involved displaying a stranger at a short (4 to 15 meters; 13 to 49 feet) or long (20 to 50 meters; 65 to 164 feet) distance in daylight for 10 seconds. While observing an overall decrease in identification accuracy as distance increased in both target present and target absent arrays, concluded that "[e]ven at 43 [meters; 141 feet], identification evidence has some diagnostic value and therefore probative value as well."[322] More importantly, they described the decline in accuracy as distance increased as "intuitively likely and consistent with previous research."[323]

---

[319] Wagenaar & Van der Schrier, *Face Recognition as a Function of Distance and Illumination: A Practical Tool for Use in the Courtroom*, Psychol., Crime & L., 321, 328, Table 2 (1996). These authors went on to create "The Rule of Fifteen: A diagnostic value of 15 is reached at not more than 15 meters, not less than 15 lux." *Id*. at 329. However, the results of this experiment do not support such a conclusion. As one court wrote, "It turned out that the so-called 'Rule of 15' was merely an idea advanced in a single article . . ." in which "it is difficult to extract . . . how the foregoing ideas were factored together to arrive at the so-called 'diagnostic value.'" *United States v. Herrera*, 788 F. Supp. 2d 1026 (NDCA 2011). It also has been discredited in the scientific literature, *e.g.*, Lampinen, Erickson, Moore & Hittson, *Effects of distance on face recognition: implications for eyewitness identification*, Psychon. Bull. Rev., 1489, 1492 (2014); Lindsay, Semmler, Weber, Brewer & Lindsay, *How Variations in Distance Affect Eyewitness Reports and Identification Accuracy*, Law & Hum. Behav., 1, 8, 9 (2008) ("the 15-m rule is not particularly useful for the courts").

[320] Wagenaar & Van der Schrier (1996), *supra*, at 325, Table 2. The study excluded those who did not make an identification. The exposure time was 3 seconds for each of 7 targets at a distance of 3 meters (10 feet). The 54 subjects were then shown 7 arrays, some target present and some target absent, each at a constant distance with increased illumination until the subject identified the target or a filler, said the target was not present, or refused to make a choice.

[321] Wagenaar & Van der Schrier (1996), *supra*, at 325.

[322] Lindsay et al. (2008), *supra*, at 8, 9. The Special Master in Henderson was under the impression that "faces are essentially unrecognizable at 134 feet." Report of the Special Master, *State v. Henderson*, at 45. This conclusion is

A third study (Kerstholt) displayed 24 photographs of co-workers (of varying levels of familiarity) and strangers for 5 seconds at a distance of 16.5 feet in a dark room with less illumination than a night with a full moon and hoods partially or fully concealing the face.[324] Under these adverse conditions, participants in a yes-no test correctly identified their co-workers at a much lower (56%) rate than they correctly rejected the strangers as unknown (78%).[325] In order to check recognition, the researchers displayed photographs of both the known and unknown faces illuminated by a flashlight, and 98.8% of the subjects correctly identified them, a rate characterized by the researchers as "nearly perfect."[326]

The most recent study (Lampinen) displayed eight different live human beings in daylight for 10 seconds each at distances ranging from 15 to 120 feet, followed by the presentation of 16 photographs (eight targets, eight fillers) in random order.[327] They found that hits decreased and false alarms increased with distance. At 120 feet, overall accuracy was about 55%. While the authors took confidence statements, the article does not reveal the relationship between confidence and accuracy at each distance or at each level of confidence. This study also does not answer the question about the ability of a person to recognize a single stranger s/he observed live for more than 10 seconds or at a distance of less than 15 feet.

Surprisingly, these four studies appear to be the only research in the social science literature on distance and/or lighting[328] with respect to the identification of strangers.[329] None involved distances of less than

---

not borne out by the 1996 Wagenaar study or the 2008 Lindsay study. Indeed, Lindsay et al. (2008), recognized that "further research would be required to determine the distance and conditions that reduce diagnosticity to one and thus probative value to zero." Lindsay et al. (2008), *supra*, at 9.

[323] Lindsay et al. (2008), *supra*, at 8.

[324] Kerstholt, Raaijmakers & Valeton, *The Effect of Expectation on the Identification of Known and Unknown Persons,* Applied Cognit. Psych., 174, 176 (1992). Exposure time, extent of head covering, and expectation also were manipulated.

[325] Kerstholt et al. (1992), *supra*, at 177, Table 1.

[326] Kerstholt et al. (1992), *supra* at 177 ("nearly perfect").

[327] Lampinen et al. (2014), *supra*, at 1489 ("This is one of the first studies to examine the ability to recognize faces of strangers at a distance under free-field conditions").

[328] "Surprisingly, we know of no experiments that have measured the light levels required for the encoding of faces." Wells & Olsen, *Eyewitness Testimony*, Ann. Rev. Psychology, 277, 282 (2003).

[329] There are three additional studies on the recognition of photographs of familiar or celebrity faces, which were manipulated to approximate different distances. Wagenaar studied the recognition of "familiar" faces at different distances and lighting levels. Overall, the proportion of accurate identifications was higher than for strangers. De Jong, Wagenaar, Wolters & Verstijnen, *Familiar Face Recognition as a Function of Distance and Illumination: A Practical Tool for Use in the Courtroom*, Psychol., Crime & L., 87, 91, Table 1 (2005). This article reported a 95% hit rate in an earlier study of celebrities at distances of 240 feet for men and 200 feet for women, which they attribute to greater illumination and a different methodology. *Id., citing* Green & Fraser, *Observation distance and recognition of photographs of celebrities' faces*, Perceptual & Motor Skills, 637 (2002). *Compare* Green & Fraser (2002), *supra*, at 646-647 (using 8 x 10" photographs, upper limit for recognition of celebrities' faces is approximately 320-340 feet) *with* G. Loftus & Harley, *Why is it easier to identify someone close than far away?*, Psychon. Bull. Rev., 43, 63 (2005) (using celebrity faces, facial identification remains constant to 25 feet and then fall offs to zero at 110 feet). This study was conducted after one of the authors had testified for the defense in a case where the witness saw the perpetrator at a distance of 450 feet. Kerstholt et al. (1992), discussed above, text at notes 24-26, was conducted for the defense in litigation to establish that a woman would not have recognized her stepfather under the test conditions. The subjects could identify their co-workers by other means if they did not know their names, which suggests that at least some of them did not know their co-workers well. "Humans are capable of astonishing performances [in familiar face recognition]; for example, they can identify individuals

10 feet. Two used simulated distances. None involved exposure times of more than 12 seconds. Only one measured different (simulated) lighting levels. Only one involved the identification of a single target. Two took confidence statements, but did not assess how accurate the subjects were at different levels of confidence; the others did not take confidence statements. And none has been replicated.

RETENTION INTERVAL, DELAY, TIME ELAPSED

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on retention interval/delay/time elapsed as follows: "Memories fade with time. As a result, delays between the commission of a crime and the time an identification is made can affect the reliability of the identification. In other words, the more time that passes, the greater the possibility that a witness's memory of a perpetrator will weaken."

Analysis

Ordinary people know from their own experience that memory may fade over time.[330] Current research indicates that the longer the time between the crime and the identification procedure, the less likely it is that a witness will make an identification. Thus, there will be *fewer* identifications.[331] However, witnesses who make high-confidence identifications after a long retention interval are *no less accurate* than those who make a high confidence identification shortly after the crime.[332] Moreover, a longer retention interval (like heightened stress and short exposure times) has "little effect on the innocent suspect false identification rate."[333] Thus, time does not increase the probability of wrongly identifying an innocent suspect.

One set of researchers observed, "The most striking feature of our examination . . . is the consistency of the findings across retention interval conditions . . . confidence and accuracy were meaningfully related for choosers in both the immediate and delayed conditions, particularly in the upper half of the

---

despite not having seen them for decades and can tell apart familiar from unfamiliar faces in a few hundred milliseconds." Ramon, Vizioli, Liu-Shuang & Rossion, *Neural microgenesis of personally familiar face recognition*, Proceedings of the National Academy of Sciences, 1 (June 30, 2015).

[330] *Bomas v. State*, 956 A.2d 215, 222 (Md. Spec. App. 2008), aff'd, 987 A.2d 98 (Md. 2010) (expert "conceded, on cross-examination, that the fact that memories dim over time was a matter that could be understood without expert testimony").

[331] Palmer et al. (2013), *supra,* 64.

[332] Palmer et al. (2013), *supra*, at 61. *See* Clifford, Havard, Memon & Gabbert, *Delay and Age Effects on Identification Accuracy and Confidence: An Investigation Using a Video Identification Parade,* Appl. Cognit. Psychol., 1, 2, Table 1 (2011) (no delay effect in 13 studies; delay effect in 11 studies).

[333] Palmer et al. (2013), *supra*, at 63-64 (citing Clark & Godfrey, *Eyewitness identification evidence and innocence risk*, Pyschon. Bull. Rev., 22, 29 (2009)). In the Palmer study, subjects viewed a live person at a distance of 10 meters (33 feet) for either 5 or 90 seconds and were asked to identify the person in an 8-person target-present or target-absent photo array (with a not-present option) either immediately or 6-8 days later. *Id*. at 58-59. In a study comparing identifications by 7/8-year-olds with 13/14-year-olds at two days and two weeks, delay negatively affected both correct identifications and correct rejections by the younger children, but it "had little or no effect on our 13/14-year-old's correct rejection rates under TA line-ups, and only a small effect upon correct identifications in TP line-ups." Clifford et al. (2011), *supra*, at 6.

confidence scale."[334] They asked: "Does an increase in retention interval undermine the meaningful CA relationships reported in recent research? These results suggest no, at least not for the retention intervals in the range used here [*e.g.*, up to 8-21 days]. For choosers in both the delayed and immediate conditions, increased confidence was associated with probable accuracy."[335]

Looking at even longer retention intervals, a 2016 study that reanalyzed data from several earlier articles, focusing on suspect identifications only, found that "high confidence suspect ID accuracy was close to 100% correct whether the retention interval was as short as 1 week or as long as 9 months."[336] The authors explained that "[i]f people learn to accurately express high confidence in a recognition decision only when their internal memory match signal is strong, and if a long retention interval weakens that signal, on average, then one would expect to find fewer expressions of high confidence as the retention interval increases. However, . . . high-confidence IDs should remain accurate, and the data suggest that they do."[337]

A study of memory for events and personal attributes concluded, "when the accuracy of memory over time was measured in terms of the percentage of recalled (rather than total possible) facts that were correct, we found that the loss function was essentially flat. That is, the accuracy of recalled events did not decay rapidly and then level off. It was the total number of correct events that declined rapidly and then leveled off."[338]

These findings are in accord with earlier research on memory of traumatic events. One studied reports of a shoot-out in the middle of a city street, where "[o]ne of the most striking results was the [witnesses'] lack of memory loss over time. Accuracy rates remained virtually unchanged 5 months after the incident," in spite of inaccurate information in the media and attempts to mislead the witnesses through biased questions. [339] Another studied memories about the sinking of a party boat with the loss

---

[334] Sauer, Brewer, Zweck & Weber, *The effect of retention interval on the confidence-accuracy relationship for eyewitness identification*, Law & Hum. Behav., 337, 334 (2010). In the Sauer study, subjects viewed a live person at a distance of 10 meters (33 feet) for 10 seconds and were asked to identify the person in an 8-person target-present or target-absent photo array (with a not-present option) either immediately or 18-21 days later. *Id*. at 341.

[335] Sauer et al. (2010), *supra*, at 344.

[336] Wixted, Read & D. Lindsay, *The Effect of Retention Interval on Eyewitness Identification Confidence-Accuracy Relationship*, J. Appl. Res. Mem. & Cog., 1, 9 (2016), reanalyzing Juslin, Olsson & Winman, *Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence-accuracy correlation*, J. Exp. Psychology, Learning, Memory & Cognition, 1304 (1996); Sauer et al. (2010), *supra*, at 334; Palmer et al. (2013), *supra*, at 61; and Read, D. Lindsay & Nichols, *The relation between confidence and accuracy in eyewitness identification studies: Is the conclusion changing?*, in Thompson et al., EYEWITNESS MEMORY: THEORETICAL AND APPLIED PERSPECTIVES, 107 (1998).

[337] Wixted et al. (2016) at 9.

[338] Ebbesen & Rienick, *Retention Interval and Eyewitness Memory for Events and Personal Identifying Attributes*, J. Appl. Psychol., 745, 760 (1998).

[339] Yuille & Cutshall, *A Case Study of Eyewitness Memory of a Crime*, J. Appl. Psychol., 291, 299 (1986) ("It appears that memory persistence results from the nature of the event, and that an Ebbinghaus decay curve simply doesn't apply in this type of case."). Other researchers have concluded from laboratory studies – over 40% of which had a null or negative effect – that "the forgetting function for the once-seen face is Ebbinghausian in nature: Rate of memory loss for an unfamiliar face is greatest right after the encounter and then levels off over time." Deffenbacher, Bornstein, McGorty & Penrod, *Forgetting the Once-Seen Face: Estimating the Strength of an Eyewitness's Memory Representation*, Exper. Pysch, 139, 148 (2008). This is one of a number of meta-analyses cited by the National Academy of Sciences as having failed to meet all current standards for conducting and reporting systematic reviews." *Identifying the Culprit* at 74, n.8 and 75-76.

of 51 lives, that found a confirmation rate over 80% between the survivors' accounts (recorded between 10 days and 2+ years later), in spite of some limitations in data collection.[340] To be sure, these studies are not of eyewitness identification but, unlike virtually all of the laboratory research, they involved traumatic events that are more likely to be remembered than everyday or neutral events.[341] Research does not support a recommendation that delays between the commission of a crime and the time an initial identification is made can affect the reliability of a high-confidence suspect identification, the kind of identification most likely to be at issue at trial. Indeed, research supports the opposite, that is, the reliability of a high-confidence identification is undiminished by delay.

DISGUISE

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on disguise as follows: "The perpetrator's use of a disguise can affect a witness's ability both to remember and identify the perpetrator. Disguises like hats, sunglasses, or masks can reduce the accuracy of an identification. Similarly, if facial features are altered between the time of the event and a later identification procedure, the accuracy of the identification may decrease.

Analysis

As is intuitively obvious, disguise appears to reduce the overall proportion of accurate identifications. But it does not eliminate accurate identifications.[342] If a person recognizes a perpetrator who is "disguised," then the disguise apparently did not work.[343] And if the disguise worked, then there is

[340] Thompson, Morton & Frazer, *Memories for the Marchioness*, Memory, 615, 629 (1997).

[341] *See, e.g.*, Yuille, Daylen, Porter & Marxsen, *Challenging the Eyewitness Expert*, in Ziskin, COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY, 1266, 1275 (5th ed. 1995)("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, 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 . . . ."); Egeth, *Expert psychological testimony about eyewitnesses: An update*, in Kessel, PSYCHOLOGY, SCIENCE AND HUMAN AFFAIRS, 151, 159 (Westview Press 1995) ("I would argue that in their zeal to disparage the ability of human memory to retain information veridically, the eyewitness experts have tended to ignore the truly counter-intuitive finding in this area. Although memory *typically* fades with time, there are circumstances in which it improves over time . . . found most often when subjects are aroused at the time of initial presentation of material. In such cases their immediate recall or recognition performance may be worse than after a couple of weeks delay. . . . [M]emory for faces may well not fit the 'typical' decreasing memory function . . . even when arousal is not an issue. That is, while some studies do show that memory for faces gets worse over time, other studies show no significant change over time while still others have found a slight improvement.") (citations omitted).

[342] Cutler, Penrod & Marten, *The Reliability of Eyewitness identification*, Law & Hum. Behav., 233, 245, Table 1 (1987) (45% of the participants identified the robber in the lineup test if he wore no hat during the robbery compared to 27% if he wore a hat during the robbery).

[343] *Baker v. United States*, 867 A.2d 988, 996 (D.C. 2005) ("Lyles testified that the person who appeared at his bedroom with the gun was wearing a ski mask, but he could still see various parts of that person, whom he later identified as Baker."); *People v. Abney*, 918 N.E.2d 486 (N.Y. 2009) (one victim immediately recognized the knife-wielding robber (whose mask left his eyes and nose exposed) as a person he had encountered regularly in the neighborhood).

unlikely to be an identification.[344] Moreover, a disguise like a hat or sunglasses appears to increase false identifications by nothing at all or not more than a percentage point or two.

"Some researchers claim that eyes are the most important facial feature for recognizing a face."[345] In one study of disguised celebrity faces, the authors observed that "Upper features of the human face convey appreciably more information for recognition than lower features . . . the eyes make more incremental contributions to recognition . . . than any other individual features."[346] In another study involving strangers, the researchers found that the students "spent just over 4 seconds of the 10 seconds of learning time examining the eyes, whereas each of the other features was examined for 1 second or less."[347] "This analysis once again shows the dominance of the eyes as an important feature for face learning."[348] A third study confirmed that "eyes are the most important facial feature for decisions related to impression formation, recognition and identification."[349] It may be, then that concealing the eyes has a greater impact on the ability of a witness to identify a stranger than other "disguises" that conceal less than the full face.

In one experiment, there was a decrease in the correct identification rate from 87% when the target wore no disguise to 78% when he wore a hat, to 69% when he wore sunglasses, to 55% when he worse both a hat and sunglasses. By contrast, the correct rejection rate hovered around 72-74% for all conditions except when the target was wearing sunglasses, when it fell to 64%.[350] Thus, while a disguise appears to reduce correct identifications, it does not substantially increase false ones. In a second experiment, there was a decrease in the correct identification rate from 80% when the target wore no disguise to 69% for a hat, to 53% for a partial stocking mask to 55% for a full stocking mask. By contrast, the correct rejection rate was 76-79% for no disguise or a hat only and 64-67% for a partial or full stocking mask.[351] Again, a disguise appears to reduce correct identifications without substantially increasing false ones – especially if the only disguise was a hat.

---

[344] *Miller v. United States*, 14 A.3d 1094, 1098-99 (D.C. 2011) ("The gunman's mask concealed his face, and neither [eyewitness] was able to identify Miller as the shooter.").

[345] Mansour, Beaudry, Bertrand, Kalmet, Melsom & Lindsay, *Impact of Disguise on Identification Decisions and Confidence with Simultaneous and Sequential Lineups*, Law & Hum. Behav., 514, (2012), *citing* Henderson, Williams & Falk, *Eye movements are functional during face learning*, Mem. & Cog., 98 (2005) and Janik, Wellens, Goldberg & Dell-Osso, *Eyes as the center of focus in the visual examination of human faces*, Perceptual & Motor Skills, 857 (2005).

[346] Fisher & Cox, *Recognizing human faces*, Applied Ergonomics, 104, 107 (1975).

[347] Henderson et al. (2005), *supra*, at 101. *See also* Janik et al. (1978), *supra,* at 858 (43.4% of subject's visual inspection time was spent looking in the region of the eyes).

[348] Janik et al. (1978), *supra,* at 858.

[349] Mansour et al. (2012), *supra,* at 522.

[350] Mansour et al. (2012), *supra,* at 518, Table 1. (Oddly, the correct rejection rate was higher for sunglasses and a hat than sunglasses alone. In real life, one of the "foils" would have been a suspect. In calculating the false identification rate (as opposed to the correct rejection rate), therefore, the "foil" identifications should be divided by six. Doing so, the false identification rate in this example would have increased from approximately 4.5% for no disguise, sunglasses and a hat, or hat only to 6% for sunglasses only. Similar results would be found in the other examples given.)

[351] Mansour et al. (2012), *supra,* at 521, Table 3.

Other studies reach conflicting conclusions on the use of a cap or hat as a disguise. One study found fewer correct judgments when a hat covering most of the hair was used,[352] whereas, another study found no such effect.[353]

Overall, confidence decreased with disguises, indicating that "witnesses are sensitive to some degree of their ability to make accurate identifications, and this sensitivity is in a form (confidence) understood by triers of fact."[354]

Again, a disguise may affect the rate of identification, but not the accuracy. A recommendation that suggests that an innocent suspect is more likely to be identified when a disguise is used is not supported by the research.

INTOXICATION

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized research on intoxication as follows: "An identification made by a witness under the influence of a high level of alcohol at the time of the incident tends to be more unreliable than an identification by a witness who drank a small amount of alcohol."

Analysis

There is little or no scientific support for the proposition that intoxication reduces identification accuracy. First, there appear to be only a few articles on the effect of intoxication on eyewitness identification performance.[355] One recent article (Hagsand, 2013) concluded, "[a]t present, it is too early

[352] Cutler, Penrod & Martens, *Improving the Reliability of Eyewitness Identification: Putting Context Into Context*, J. Appl. Psych., 629, 633 & 634, Table 3 (1987) (Mean accuracy of 51% for no hat vs. 40% for hat).

[353] Yarmey, *Eyewitness Recall and Photo ID: A Field Experiment,* Psychology Crime & Law, 65 (2004) ("Target disguise [baseball cap and dark sunglasses] . . . had no significant main effects on identification in either the target-present or the target absent lineups").

[354] Mansour et al. (2012), *supra,* at 524.

[355] "To our knowledge, there are only two published studies on the effects of alcohol on eyewitness lineup performance." Hagsand, Roos-af-Hjelmsater, Granhag, Fahlke & Soderpalm-Gordh, *Do Sober Eyewitnesses Outperform Alcohol Intoxicated Eyewitnesses in a Lineup*, The European J. Psychology Applied to Legal Context, 23, 25 (2013) (Hagsand I), *citing* Yuille & Tollestrup, *Some effects of alcohol on eyewitness memory*, J. Applied Psych, 268 (1990); and Dysart, Lindsay, MacDonald & Wicke, *The intoxicated witness: Effects of alcohol on identification accuracy from showups*, J. Appl. Psych., 170 (2002)). The Hagsand article makes three. One more article comparing intoxication and cross race effects is discussed below. There are many studies of the effect of alcohol on memory generally. One review found that "On recognition memory tasks, alcohol has been shown to decrease hit rates . . . but to have no effect on false alarm rates." Mintzer, *The acute effects of alcohol on memory: A review of laboratory studies in healthy adults*, Int. J. Disabil. Hum. Dev., 397, 399 (2007) (parentheticals omitted). Another study of eyewitness recall, not identification, found that although there were fewer details in the alcohol condition, "there was no difference in accuracy rate as a function of alcohol dose." Hagsand, Roos-af-Hjelmsater, Granhag, Fahlke & Soderpalm-Gordh, *Bottled memories: On how alcohol affects eyewitness recall*, Scandanavian J. Psych., 188, 193 (2013) (Hagsand II).

to draw conclusions about the impact of alcohol intoxication on eyewitnesses' identification ability, thus more studies in this field are recommended."[356]

A 2016 article found that "intoxicated participants were no less likely than sober or placebo participants to make an accurate identification from a TP [target-present] lineup," and there was "no significant association between alcohol condition and identification decision" in a TA [target-absent] lineup."[357]

It appears that alcohol has little or no negative effect on eyewitness performance. The Hagsand study found that "intoxicated eyewitnesses performed on the same level as their sober counterparts."[358] The results actually could be interpreted to suggest that alcohol increases eyewitness performance. In target present lineups, 40% in the higher alcohol dose group correctly identified the target compared to 5% in the lower alcohol dose group and 25% in the no alcohol control group. Similarly, in target absent lineups, 45% of the higher alcohol dose group correctly rejected the lineup compared to 36.4% in the lower alcohol dose group and 23.8% in the no alcohol control group.[359]

Hagsand cited two articles. One stated "alcohol intoxication did not have an effect on identification accuracy in either the TP condition or the TA condition." The other concluded, "the level of intoxication did not have a significant effect on performance in the TP lineup, but as the level of intoxication increased, participants in the TA lineup made fewer rejections, which gives support for the alcohol myopia theory."[360]
An experiment that compared a one-photo "show-up" between persons with a blood alcohol level of .04 or less (low alcohol) and above .04 (high alcohol) found that there was no difference in identification accuracy in a target-present photo show-up, but a higher rate of false identifications in target-absent photo show-ups in the high alcohol condition,[361] except when instructions were given to the witness to be cautious – which resulted in a 97% accuracy rate for intoxicated witnesses in target-absent show-ups.[362]

---

[356] Hagsand I, *supra*, at 42.
[357] Kneller & Harvey, *Lineup Identification accuracy: The effects of alcohol, target presence, confidence ratings, and response time*, European J. Psychology Applied to Legal Context, 11, 19 (2016). In fact, 45% of the subjects in the alcohol condition rejected the lineup compared to 50% in the placebo condition and 40% in the control (no-alcohol) condition.
[358] Hagsand I, *supra*, at 23 ("[E]yewitnesses who have consumed a lower (0.4 g/kg ethanol) or a higher (0.7 g/kg ethanol) dose of alcohol perform at the same level as sober eyewitnesses in a lineup.").
[359] Hagsand I, *supra*, at 37, Table 1 and 38, Table 2.
[360] Yuille & Tollestrup (1990), *supra,* and Dysart et al. (2002), *supra*, respectively. A study of the effects of alcohol on a *thief* caught in the act, found that "[i]n the presence of higher arousal [*e.g.*, stress], the alcohol subjects performed as well as the placebo subjects, suggesting that these subjects overcame the detrimental effects of alcohol consumption." Read, Yuille & Tollestrup, *Recollections of a Robbery: Effects of Arousal and Alcohol upon Recall and Person Identification*," Law & Hum. Behav., 425, 442 (1992).
[361] Dysart, Lindsay, MacDonald & Wicke, *The Intoxicated Witness: Effects of Alcohol on Identification Accuracy From Show-ups*, J. Appl. Psych., 170, 173 (2002).
[362] Dysart & Lindsay, *Intoxicated witnesses: The effect of clothing and instruction bias on identification accuracy from show-ups* (under review 2006), *cited in* Dysart & Lindsay, *Show-up Identifications: Suggestive Technique or Reliable Method*, in Lindsay, Ross, Read & Toglia, HANDBOOK OF EYEWITNESS PSYCHOLOGY, Vol. II,137, 149 (2007). This article concludes, "Combined, the results of these experiments suggest that intoxicated witnesses presented with show-ups can be as accurate, or more accurate, than sober witnesses under certain circumstances." *Id*. "Show-up" is in quote marks in the text because the laboratory studies use a single photograph and not a real person. It is difficult to assess whether the results would be the same if a real person was shown to the witness.

One study compared both "intoxication" and cross-race identification. Those in the alcohol condition made fewer identifications – but alcohol diminished the difference between accurate same- and cross-race identifications to 2 percentage points (81% v. 79%, "a small tendency,") compared to 5% in the non-alcohol condition (86% v. 81%, "significantly higher"). This study is notable in that the entire span of accurate identifications regardless of alcohol consumption or racial differences was seven percentage points (79% to 86%).[363] There may be a statistical difference between the alcohol and non-alcohol conditions, but there does not appear to be a practical difference.

It may be that with higher levels of intoxication than those studied, overall identification accuracy would decrease. But even if there were research on that point, ordinary citizens are well qualified to make an assessment on to the effect of intoxication as anyone else and, based on the science that exists, it would be inappropriate to make a recommendation that suggests that intoxicated witnesses are "unreliable."

PRIOR DESCRIPTION

Research as summarized in the *Henderson* Jury Instruction

The *Henderson* jury instruction summarized the research on prior description as follows: "Another factor for your consideration is the accuracy of any description the witness gave after observing the incident and before identifying the perpetrator. Facts that may be relevant to this factor include whether the prior description matched the photo or person picked out later, whether the prior description provided details or was just general in nature, and whether the witness's testimony at trial was consistent with, or different from, his/her prior description.

Analysis

Although similarities or differences between the description of the perpetrator and the defendant will undoubtedly be argued by one of the parties to either bolster or diminish the strength of the identification, scientific support for it is limited. Indeed, the research points in the other direction. A 2014 review of the literature on this subject found that "despite the clear intuition that witnesses who are better at describing a target should also be better at recognizing it, this relationship has proved to be quite elusive and generally weak.[364] Moreover, "efforts to describe a previously seen face can actually *impair* subsequent memory performance, at least under some circumstances."[365] Although, there is no evidence of its effect on false identifications.[366]

There apparently are exceptions:

---

[363] Hilliar & Kemp, *Now Everyone Looks the Same: Alcohol Intoxication Reduces the Own-Race Bias in Face Recognition*, Law & Hum. Behav., 367, 372, 373, Figure 2 (2010).

[364] Meissner, Sporer & Schooler, *Person Descriptions as Eyewitness Evidence*, in Lindsay, Ross, Read & Toglia, eds. THE HANDBOOK OF EYEWITNESS PSYCHOLOGY, Vol. II, 3, 21 (2014).

[365] Meissner et al. (2014), *supra*, at 22; *id*. at 21 ("the elicitation of elaborate verbal descriptions may lead participants to generate inaccurate details which then impairs their recognition performance*")*.

[366] Mickes & Wixted, *On the Applied implications of the "Verbal Overshadowing Effect,*" Perspectives on Psychol. Science, 400, 401 (2015) ("in both the original study (Schooler) and the replication studies (Alonga) [of the verbal overshadowing effect], memory was tested using only target-present lineups (i.e., lineups containing the previously-seen target face), making it possible to compute the correct ID rate (i.e., the hit rate) but not the false ID rate (i.e., the false alarm rate).").

> [1] inasmuch as the recognition of other-race faces depends on the quality of witnesses' memory of individual features, the veracity of witnesses' memory for those features (as revealed by the quality of their descriptions) becomes predictive of their recognition performance. . . . and
>
> [2] distinctive faces tended to be easier to describe and to recognize than less distinct faces, thereby leading to a modest relationship between recognition accuracy and description quality across faces. . . .[367]

The authors of this review caution that:

> Whereas in general it is useful for witnesses to generate as much information about a witnessed event as possible, in the context of person description, encouraging people to spend extensive time generating their descriptions can actually impair face recognition and result in the generation of a greater proportion of inaccurate details.[368]

In light of the considerable research on this subject that disclaims a relationship between description accuracy and identification accuracy, a recommendation similar to that in the *Henderson* instruction would be affirmatively misleading.

[367] Meissner et al. (2014), *supra*, at 21.

[368] Meissner et al. (2014), *supra*, at 26. *Accord* Meissner, *Applied Aspects of the Instructional Bias Effect in Verbal Overshadowing*, Appl. Cog. Psychol., 911 (2002) ("Prompting or encouraging a witness to reach beyond their criterion of initial recall may promote rather deleterious effects both on subsequent attempts at recall and eventual requests for perceptual identification.").

CONCLUSION

The absence of support in the social science research and, more importantly, the presence of inconsistent or contrary results suggest that the Courts should not embrace research that has been given prominence in both the media and in litigation. A closer look at some of the old research indicates that it is not as strong or as universal as it has been portrayed. As many of these studies find, a reduction in overall accuracy does not necessarily mean that witnesses are more likely to falsely identify an innocent suspect. A reduction of overall accuracy by a few percentage points does not necessarily mean that witnesses in one condition are less accurate than in another.

Only suspects are put at risk when they are identified. Suspects who are not identified, either because the witness picks a filler or does not pick anyone at all, are not at risk. Recent research that evaluates (or re-evaluates earlier data on) suspect identifications – whether true or false – has concluded that witnesses adjust their confidence to account for conditions, and that high confidence identifications (90-100%) are highly accurate, that is, in the 95-100% range,[369] regardless of factors that may decrease the proportion of high confidence identifications or any identification at all.

[369] Wixted, Mickes, Clark, Gronlund & Roediger III, *Initial Eyewitness Confidence Reliably Predicts Identification Accuracy*, Am. Psychol. Assn., 515, 520 (September 2015).

GLOSSARY OF TERMS USED IN EWID LITERATURE

Archival study – A retrospective analysis of data obtained from real police cases.

Base rate – The percentage of targets/perpetrators in the set of arrays presented to subjects/witnesses. Laboratory-based studies that examine both Target Present [TP] and Target Absent [TA] arrays usually have a base rate of 50% (half TP and half TA). There is reason to believe that in cases where there must be at least reasonable suspicion to put a suspect into an array, the base rate of guilty suspect lineups would be higher and the probability of identifying an innocent suspect would be lower.[370] Thus, the false identification rates obtained in the laboratory are likely inflated for jurisdictions that require reasonable suspicion for placing a suspect in a lineup.

Biased instructions – "Biased lineup instructions either fail to warn the witness that the culprit might not be in the lineup or imply that the culprit is in the lineup."[371] (See Unbiased instructions.)

Blind (or double-blind) administration — In a blind (or double-blind) procedure, an individual who does not know the identity of the suspect or the suspect's position in the photo array shows a photo array to the eyewitness. NAS at 24.[372]

Blinded administration – In a blinded procedure, an individual who does know the identity of the suspect shows a photo array to the eyewitness but is unable to tell when the witness is looking at the suspect's photo . . . ." NAS at 24-25. (See Blind or double-blind.)

Discriminability – The ability to tell the difference between innocent and guilty suspects. (See Response bias.) An identification procedure with high discriminability allows eyewitnesses to distinguish between suspects who are innocent vs. suspects who are guilty in the same way that a medical test with high discriminability allows doctors to distinguish between patients who have a disease vs. patients who do not.

---

[370] Gary Wells, who once described confidence as "forensically useless," recently acknowledged that "calibration tends to be good for those making an identification." Reanalyzing some of his own data, he found "nearly perfect calibration for very high confidence witnesses," when the base rate was 70%. Wells, Yang & Smalarz, *Eyewitness Identification: Bayesian Information Gain, Base-Rate Effect-Equivalency Curves, and Reasonable Suspicion,* Law & Hum. Behav., 99, 118-119 (2015). Wells advocated for having at least a reasonable suspicion to put a suspect in an array. *Id.* at 117 ("For a police department running a 75% base rate, 91 of every 100 witnesses who identify the suspect will be accurate and 9 of every 100 identifications of the suspect will be cases of mistaken identification").

An unpublished archival study found a base rate of 95% when DNA established that the perpetrators absolutely were or were not in the arrays. Kellstrand, *Eyewitness identification accuracy in cases accepted and rejected for prosecution: An archival analysis of criminal case files*, Unpublished Manuscript, San Diego, University of California (2006) (*cited in* Clark & Wells, *On the Diagnosticity of Multiple-Witness Identification*s, Law & Hum. Behav., 406, 416 (2008)). Counting cases where DNA was inconclusive, the base rate was still 68%. *See also* Behrman & Davey, *Eyewitness identification in actual criminal cases: An archival analysis*, Law & Hum. Behav., 475 (2001) (estimating an 80% base rate), *discussed in* Malpass, *A Policy Evaluation of Simultaneous and Sequential Lineups*, Psychology, Public Policy, and Law 394, 401-402, 404 (2006) (exploring the entire range of possible values).

[371] Wells et al. (2015), *supra*, at 109. We are not aware of research on failure to warn.

[372] The term "double-blind" is somewhat of a misnomer because the witness may know who the perpetrator is even if the investigator does not.

Effect size – The actual size of the difference between two (or more) groups with respect to the variable being tested. This measure provides information about the importance of a result because effect sizes can be very small (and thus not very important) despite being statistically significant.

Estimator variables – Conditions present during memory formation or storage, that are beyond the control of the criminal justice system (*e.g.*, distance, stress, race).

Exposure time, exposure duration – The amount of time that a target/perpetrator is exposed to the subject/witness. In mock scenarios, it is not always clear whether the term refers to the entire length of the video or skit, or to the amount of time the target's face is visible.

Facial recognition study – An experiment in which subjects are shown a series of photographs in succession (*e.g.*, one photograph every 2 seconds) and then are asked to indicate which faces are "old" or "new" from a larger or different set of photographs, some of which appeared in the first series (old) and some of which did not (new). For example, in a face-recognition study on the cross-race effect, the subjects might be shown a set of 5 same-race faces and 5 cross-race faces and then tested with a set of 10 same-race faces and 10 cross-race faces, half of which were in the first set.

False identification – The identification of an innocent suspect as the perpetrator. The term is used inconsistently in the literature and may refer to various incorrect decisions (i.e., identifications of fillers). However, it is only the selection of an innocent suspect that can lead to a false conviction.

Field study – Either (1) a contemporaneous study of real police cases (*e.g.*, Houston Police Department field study), or (2) a study conducted in public spaces (*e.g.*, Palmer et al. (2013), Study 1).

Forensically relevant study – An experiment in which subjects are shown only one target each and tested on only one lineup or show-up to more closely replicate an eyewitness's experience.

Filler or foil – A person or photograph in a lineup or photo array who are known to be innocent. The selection of a filler or foil cannot lead to a false conviction.

Ground truth – A term used in many fields (starting with geology) to describe information provided to decision makers from their own direct observation (direct evidence) as opposed to information provided by inference (circumstantial evidence). In laboratory experiments, researchers (the decision makers in a lab study) know the identity of the target (i.e., they know ground truth because they themselves placed the target in the lineup) and, therefore, know whether the subjects correctly identified the target; in real life, investigators (the decision makers in the early stages of a criminal investigation) do not know the identity of the perpetrator (i.e., they do not know ground truth) and must therefore try to determine whether witnesses correctly identified a stranger based on the circumstances of the crime/observation, the identification procedures, and, if available, corroborating evidence.

High confidence – Expressing 80-100% or 90-100% certainty, or using words like "Sure," "Positive," "Absolutely certain," "Will never forget that face," "Positive he did it," "I am sure that it is him," "Very sure," "Definitely him," "Looks exactly like," and "That's him! That's the guy!"

Lineup – In eyewitness identification research, a photo array typically composed of the suspect's photograph and the photographs of five fillers/foils in the US, or the suspect's video and the videos of

eight fillers/foils in the UK (not a live lineup). In real life, the presentation of a suspect (not a photograph) with five other persons as fillers/foils.

Low confidence – Expressing 60% or less certainty or using words like "Looks similar, "Possibly the guy," "Maybe," "I think but I am not sure," and "Not too sure."

Medium confidence – Expressing 70-80% or 70-90% certainty or using words like "Very similar," "Looks most like," "Pretty sure," "Fairly sure," "Looks like," "Almost certain," and "Almost a perfect match."

Meta-analysis – A method for statistically combining the results across a series of studies that meet pre-established criteria for inclusion. The National Academy of Science found fault with the meta-analyses on which many conclusions about eyewitness identification reliability have been based, saying "[n]one of the reviews met all current standards for conducting and reporting systematic reviews, and few met even a majority of these standards, making assessment of the credibility of their findings problematic. After examining the reviews, the committee concluded that the findings may be subject to unintended biases and that the conclusions are less credible than was hoped."[373]

Perpetrator/Offender/Culprit – The person who committed the crime. (See Target.)

Reliable – Dependable; trustworthy. For example, in eyewitness identification, a person who says s/he is 90-100% confident is correct 95-100% of the time. In other words, high-confidence identifications are reliable (i.e., they are trustworthy). [374]

Response bias – The willingness to make an ID. (See Discriminability.) A liberal response bias refers to a strong tendency to make an ID, which results in the benefit of a high rate of correctly identifying guilty suspects, but that benefit comes at the cost of a high rate of misidentifying innocent suspects. A conservative response bias refers to a general reluctance to make an ID, which results in the benefit of a low rate of misidentifying innocent suspects, but that benefit comes at the cost of a low rate of correctly identifying guilty suspects.

Retention interval, time elapsed, delay – The amount of time between first seeing the target/ perpetrator and the recognition test (photo array, lineup, show-up). In the social science literature, it ranges from immediately to two years.

Show-up – In laboratory research, the display of a single *photograph* (typically conducted immediately to a week after the exposure.[375] In field studies and in real life, the display of a single human being shortly (within two hours) after the initial exposure or crime.

---

[373] National Academy of Sciences, Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION, 75-76 (2014)

[374] The witnesses in the DNA exoneration cases also were reliable. None of them initially identified the suspect with high confidence, thereby signaling that they were not sure. *See* Garrett, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG (Harvard U. Press 2011). Had the police recognized the significance of their uncertainty, these innocent suspects likely would not have been arrested and prosecuted.

[375] The display of a single photograph of a stranger is not permitted in real life even if it was done close in time (within 2 hours) of the crime.

Statistical significance – "In normal English, 'significant' means important, while in statistics 'significant' means probably true (not due to chance). A research finding may be true without being important." www.surveysystem.com/signif.htm. "Something can be statistically significant and yet have a small effect size."[376] For example, if a new procedure for weight loss results in a statistically significant difference, the "not by chance" finding may reflect a difference of an ounce or two, an amount that has little practical significance. *See* Effect size.

Statistical methods

(1) Calibration – A method of assessing accuracy at various levels of confidence. Accuracy here refers to the accuracy of the set of witnesses who identify suspects and fillers. Thus, the accuracy reported in a calibration study does not directly apply to the set of witnesses whose accuracy we are primarily interested in measuring. The set of witnesses whose accuracy we are primarily interested in measuring are the ones who ID a suspect (not a filler) because that is the set of witnesses who testify in court and for whom the question of accuracy arises. *See* Confidence-Accuracy Characteristic, *infra*. The formula for determining the calibration at different levels of confidence is:

$$\frac{\text{Correct IDs in TP arrays}}{\text{Correct IDs in TP arrays + False IDs in TA arrays + Filler IDs in TA arrays (+ Filler IDs in TP arrays)}}$$

In studies that measure calibration, Filler IDs in TP arrays are not always included in the computation.

(2) Perfect calibration – A person who says he/she is 100% confident is correct 100% of the time, and so forth for each level of confidence, *e.g.*, a person who says s/he is 60% confident is correct 60% of the time. When fillers are included in the analysis (in which case, the results are not directly relevant to the question at hand), witnesses are often found to be over-confident. However, when fillers are not included in the analysis (in which case the results are directly relevant to the question at hand), recent laboratory studies show that people are under-confident rather than over-confident. Thus, people who are 90-100% confident are correct about 98% of the time; people who are 70-80% confident are right about 94% of the time, and people who are 0-60% confident are right about 83% of the time.[377]

(3) Receiver Operating Characteristic/ROC – A plot of correct suspect ID rates to false suspect ID rates for different levels of response bias (defined above). It is used to measure which procedure (for example, sequential or simultaneous presentation) yields better discriminability (defined above). Using three levels of confidence, for example, the number of high-confidence correct suspect IDs ÷ total number of TP arrays is calculated (this is the high-confidence correct ID rate); the same is done for the high-confidence false suspect ID, dividing the denominator by the number of faces in the array when there is no designated suspect (this is the high-confidence false ID rate). Next, the high- and medium-confidence correct and false ID rates are counted and divided by the total number of TP and TA arrays, respectively (again correcting the latter for lineup size if there is no designated innocent suspect), yielding a second pair of correct

[376] Wells, *Scientific Status*, in Faigman, Kaye, Saks & Sanders, MODERN SCIENTIFIC EVIDENCE, 475, 496 (2005).

[377] Wixted, Mickes, Clark, Gronlund & Roediger III, *Initial Eyewitness Confidence Reliably Predicts Identification Accuracy*, American Psychological Assn., 515, 519 (September 2015) ("[M]ost would probably agree that the 17% error rate is too high to justify a conviction based on a low-confidence ID alone.").

and false ID rates. Finally, the high- plus medium- plus low-confidence correct and false ID rates are computed in the same way. This last ROC point is the overall correct and false ID rates that are usually reported in non-ROC studies of eyewitness identification. The results for the three correct ID rates are plotted on the vertical axis and those for false ID rates are plotted on the horizontal axis, starting with high confidence only, then high plus medium confidence, then high plus medium plus low confidence. The procedure that yields a ROC curve that falls farther from the diagonal line of chance performance indicates that it better enables eyewitness to correctly sort innocent and guilty suspects into their correct categories (*e.g.*, it is the procedure that yields higher discriminability).

(4) Confidence-Accuracy Characteristic/CAC – A statistical method used in eyewitness identification studies to ascertain the reliability of suspect identifications by determining the eyewitnesses' accuracy at each level of confidence. It is just like calibration except that it includes only suspect IDs (no filler IDs). The formula is:

$$\frac{\text{Total correct suspect IDs in TP arrays}}{\text{Total correct suspect IDs in TP arrays + Total false suspect IDs in TA arrays (divided by the number in the array if there is no designated false suspect)}}$$

The calculation is performed separately for IDs made at each level of confidence. For example, in a study in which there is no designated innocent suspect in TA arrays, if there were 25 correct suspect IDs made with high confidence in TP arrays and 2 incorrect filler IDs made with high confidence in TA arrays with 6 lineup members, then high-confidence accuracy would be: $25 \div [25 + (2 \div 6)] = 98.7\%$.

(5) Point bi-serial correlation coefficient, "r" – An analysis that is related to calibration analyses and CAC analyses in that it provides a measure of the relationship between confidence and accuracy. However, instead of plotting the relationship between confidence and accuracy, the relationship is boiled down to a single number between 0 and 1. Zero means no correlation, 1 means perfect correlation; positive values means when one goes up so does the other, negative values mean the when one goes up the other goes down. The measure is based on the aggregate of true and false suspect and filler identification made with different levels of confidence. There is a complicated formula that boils the data down to a single number. Expert witnesses usually testify that r = .41, and characterize it as a modest or moderate relationship between confidence and accuracy.[378] Anything over .5 is considered to be a strong relationship

[378] By way of comparison, one of the first studies on AZT and HIV show results that were considered to be so strong that an independent board stopped the experiment on the grounds that it was unethical to withhold the drug from people in the placebo control group. In that study the effect size was .23. This formula can be – and has been – misinterpreted. The "point-biserial correlation may underestimate, or even hide, a useful relation between subjective and objective probabilities of correct identification." Juslin, Olsson & Winman, *Calibration and Diagnosticity of Confidence in EWID: Comment on What Can Be Inferred From the Low Confidence-Accuracy Correlation*, J. Exp. Psychol., 1304, 1305 (1996). Although a high point-biserial correlation coefficient . . . indicates a strong relationship between confidence and accuracy, "a low point-biserial correlation coefficient . . . does not necessarily indicate a weak relationship." Wixted, Mickes, Clark, Gronlund & Roediger III, *Initial Eyewitness Confidence Reliably Predicts Identification Accuracy*, American Psychological Assn., 515, 517 (September 2015). For example, the Salk vaccine saved over a million lives and eventually eradicated polio, yet the experimental trials showed r = .01.

between confidence and accuracy. However, this method does not inform the police, prosecutors, the court, or juries how accurate (and hence how reliable) a witness is at a given level of confidence.[379] Calibration or, especially CAC analyses, do that.

(6) Diagnosticity ratio – The correct suspect ID rate divided by the false suspect ID rate. It is usually computed without regard for confidence, and it indicates the odds that an ID is correct.[380] It can also be computed separately for each level of confidence to show, for example, that the diagnosticity ratio increases as confidence increases. It is a useful measure for that purpose, but it is not useful for comparing the diagnostic accuracy of different eyewitness identification procedure because it conflates discriminability and response bias. It also not as useful as a CAC in conveying the significance of a level of confidence (*e.g.*, identifications made with 90-100% confidence are 95-100% accurate). According to the National Research Council (2014) report, ROC analysis is a better way to compare eyewitness identification procedures.

Subject or participant – The "witness" in laboratory studies.

System Variables – The procedures and practices law enforcement uses to elicit eyewitness identifications (*e.g.*, blind administration, witness instructions, simultaneous v. sequential presentation).

Suspect – The person the police have identified as a possible perpetrator.

Target – The "perpetrator" or photograph(s) to be identified in laboratory studies.

Target Present/TP –The target/perpetrator is in the array.

Target Absent/TA – The target/perpetrator is not in the array. Note that in some TA studies there is a designated innocent target who is chosen because of his resemblance to the true target, something the police cannot do except by chance. (In a few of the DNA exoneration cases, it was noted that the innocent suspect bore a "striking resemblance" to the true perpetrator, but this appears to be rare.)

Unbiased instructions – "Unbiased instructions . . . warn the witness that the culprit might not be in the lineup."[381]

---

[379] Some studies without a designated suspect have failed to divide the filler IDs by the number in the array. This artificially inflates the rate of false identifications because, in real life there would be only one suspect in an array and the others would be fillers. For example, an experiment may report that 60% of subjects who were shown a target absent lineup falsely identified an innocent person. However, in a fair six-person lineup only one-sixth of those identifications (10%) would be false identifications of an innocent suspect, and the other five-sixths would be filler identifications.

[380] For example, if the correct identification rate is .50 and the false identification rate is .10, the diagnosticity ratio would be 5.0.

[381] Wells et al. (2015), *supra*, note 2. The term "unbiased instructions" is a misnomer. Such instructions induce more conservative responses, that is, they reduce the number of identifications compared to a condition in which they are encouraged to make an ID. However, they do not affect overall accuracy (i.e., different instructions do not affect the accuracy of identifications and non-identifications combined). *See* Clark, Moreland & Gronlund, *Evolution of the empirical and theoretical foundations of eyewitness identification reform,* Psychon. Bull. Rev., 251 (2013).