NATIONAL POLICING INSTITUTE
Pursuing Excellence Through Science and Innovation

2025 Nationwide Survey of Eyewitness Identification Practices among Law Enforcement Agencies

Karen L. Amendola, Ph.D.
Maria Valdovinos Olson, Ph.D.
Curt Carlson, Ph.D.
Scott Grunlund, Ph.D.
Laura Mickes, Ph.D.
Jie Gao, M.S.
Yukun Yang

2025

Nationwide Survey of Policies and Practices in Eyewitness Identification: Full Report
Final Research Report to the National Institute of Justice

Final Revision December 31, 2025

Award #: 2019-75-CX-0009

"Confidence, Latency, and Accuracy in Eyewitness Identifications Made from Showups: Evidence from the Lab, the Field, and Current Law Enforcement Practices"

Karen L. Amendola, Ph.D., *Principal Investigator*
Chief Behavioral Scientist
National Policing Institute
2550 S. Clark St., Suite 1130
Arlington, VA. 22202
kamendola@policinginstitute.org
(202) 833-1460 office
(571) 562-1615 mobile
ORC ID: 0000-0002-7649-5236

Award period: January 1, 2020 – December 31, 2024

Grant Amount: $662,407 (across two studies and a national survey)

This project was supported by Award No. 2019-75-CX-0009, awarded by the National Institute of Justice (NIJ), Office of Justice Programs, U.S. Department of Justice to the National Policing Institute. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice or the National Policing Institute.

This report serves as a follow-up to the 2013 report by the Police Executive Research Forum (PERF) also funded by the NIJ, Grant No. 2010-IJ-CX0032.

Cite as: Amendola, K. L., Valdovinos Olson, M., Carlson, C., Gronlund, S., Mickes, L., Gao, J., and Yang, Y. (2025). Confidence, Latency, and Accuracy in Eyewitness Identifications Made from Showups: Nationwide Survey of Policies and Practices in Eyewitness Identification, Final Report by the National Policing Institute to the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice.

National Policing Institute
2550 S. Clark Street, Suite 1130
Arlington, VA 22202
(202) 833-1460
www.policinginstitute.org

FINAL RESEARCH REPORT:

Confidence, Latency, and Accuracy in Eyewitness Identifications Made from Showups:
Nationwide Survey of Policies and Practices in Eyewitness Identification

By:

Karen L. Amendola, Ph.D.
Maria Valdovinos Olson, Ph.D.
Curt Carlson, Ph.D.
Scott Gronlund, Ph.D.
Laura Mickes, Ph.D.
Jie Gao, M.S.
Yukun Yang, Ph.D. student

Acknowledgments

The authors wish to acknowledge the following individuals for their contributions to the nationwide survey content. First, the authors thank Patricia A. Riley, retired Special Counsel to the Acting U.S. Attorney, and former Assistant USA, now a professor with Georgetown Law and Robert "Rocky" Kravetz, former Assistant U.S. Attorney and Chief of Appeals for the District of Delaware, now partner at Bernstein, Litowitz, Berger & Grossman LLP.

We also acknowledge former NPI staff members Travis Riddle, Ph.D. and Travis Taniguchi, Ph.D. for conducting the stratified sampling and providing other assistance. We sincerely appreciate the support and assistance provided by the NPI leadership team, and especially Colby Dolly, Ph.D., Director, for his ongoing support.

In addition, we appreciate the previous efforts of the Police Executive Research Forum (PERF) for their comprehensive survey and report published in 2013, which served as a baseline and model for this survey.

We are also grateful to two individuals at NIJ: Dr. Brett Chapman, Ph.D., Senior Social Science Analyst for his input and guidance across the entire study, and Cathy Girouard, Senior Grants Management Specialist, who provided important assistance as our Grant Manager for this project.

Above all, we thank the numerous agencies that participated in this survey.

Table of Contents

List of Tables

About the Authors

Karen L. Amendola, Ph.D. (Principal Investigator or PI)

Karen L. Amendola, *Chief Behavioral Scientist,* National Policing Institute, earned her Ph.D. and M.A. in Psychology (Industrial/Organizational) from George Mason University (VA) and an M.A. in Human Resources Management from Webster University (MO). Amendola serves as the PI on an NIJ-funded study of organizational stress and an Arnold Foundation-funded experiment on supervision/procedural justice. Focusing on applied psychology in policing, she has been awarded, directed and/or overseen research, assessment, and training contracts and grants in excess of $12 million (federal, state, local, and private) in the areas of culture, early warning systems, ethics, hiring/promotion, shift scheduling, wellness, and police-community relations. In 2020–2024, Amendola served on APA's Presidential Task Force on Use of Force, and in 2021 on the Committee on *Psychological Screening for Implicit and Explicit Bias* for the California POST. In 2018–2019, she served as Chair, *Division of Experimental Criminology,* American Society of Criminology. Her publications appear in *Criminology & Public Policy,* the Journals *of Experimental & Quantitative Criminology,* among others. Amendola was also Associate Editor (Psychology and Law) for the Encyclopedia of Criminology and Criminal Justice (Bruinsma and Weisburd, Eds., 2014).

ORC ID: 0000-0002-7649-5236

Maria Valdovinos Olson, Ph.D.

Maria Valdovinos Olson, *Senior Research Associate,* National Policing Institute, earned her Ph.D. in Sociology from George Mason University. An experienced researcher and project manager, she works on a diverse portfolio of federally and privately funded research projects in policing, corrections, and prisoner reentry. She has published in the *Journal of Offender Rehabilitation*; *Policing: An International Journal of Police Strategies & Management;* and *Women & Criminal Justice*, among others, and authored a chapter in the Handbook on Prisons and Jails (2023). Her work also includes reports for the U.S. Department of Justice.

ORC ID: 0000-0002-5927-7296

Curt Carlson, Ph.D.

Curt Carlson, Ph.D., *Professor of Psychology,* East Texas A&M University, earned his Ph.D. from the University of Oklahoma in 2008, where his research involves recognition memory and face processing, particularly focusing on eyewitness memory. This research has been published recently in *Applied Cognitive Psychology*; *Psychology, Crime, and Law;Journal of Applied Research in Memory and Cognition*; and *Cognitive Research: Principles and Implications*. He serves on the editorial board for *Psychology, Crime, and Law* and the *Journal of Applied Research in Memory and Cognition*; and is a fellow of the Psychonomic Society. Dr. Carlson's research has been supported by grants from the NIJ and the American Psychology-Law Society.

ORC ID: 0000-0002-3909-7773

Scott Gronlund, Ph.D. (Co-PI)

Scott Gronlund, Ph.D., *Professor Emeritus of Psychology*, University of Oklahoma was the Roger and Sherry Teigen Presidential Professor. He earned his Ph.D. from Indiana University and

conducted research on recognition memory for over 35 years, the last 15 of which focused on eyewitness research. Professor Gronlund published two review chapters on the Psychology of Learning and Motivation (2015, 2018) on system variable reforms and on the new science of eyewitness memory. He has published in the *Journal of Experimental Psychology: Learning, Memory, and Cognition*; and the *Journal of Applied Research in Memory and Cognition*, among others. Gronlund's eyewitness research was supported by grants from the National Science Foundation. Professor Gronlund served on the Editorial Board of *Psychology, Crime, and Law*, and was a Consulting Editor for *Cognitive Research: Principles and Implications*. He is a Fellow of the Association of Psychological Science (APS) and the Psychonomic Society.

Laura Mickes, Ph.D. (Co-PI)

Laura Mickes, *Professor of Psychology,* School of Psychological Science, University of Bristol earned her Ph.D. from the University of California-San Diego, and has been researching recognition memory for 20 years. In the last decade, she has focused on eyewitness research. Professor Mickes recently published work in *Cognitive Research: Principles and Implications* (using artificial intelligence in eyewitness memory), *Behavior Research Methods* (developed a toolkit to analyze/model eyewitness ID data), *Journal of Applied Research in Memory and Cognition* (confidence-accuracy characteristic analysis), *Journal of Experimental Psychology: Applied (*diagnosticity in sequential v. simultaneous lineups)*, and Psychonomic Bulletin and Review* (sleep studies). Mickes has received grants from the National Science Foundation, the U.K. Economic & Social Research Council, and the British Academy. Professor Mickes is the Psychonomic Society's Digital Content Editor and Consulting Editor for *Cognitive Research: Principles and Implications*, and a Fellow of the Psychonomic Society.

ORC ID: 0000-0002-8090-9753

Jie Gao, M.S.

Jie Gao, *Research Data Scientist*, National Policing Institute, earned her M.S. in Statistics with a focus on machine learning and big data from the University of Washington, and her B.S. in Mathematics & Mathematical Decision Sciences from the University of North Carolina-Chapel Hill. Gao plays a critical role in research, assistance, and innovation projects and manages the NPI's data science program. She serves as a resource to team members across NPI divisions/partners to find solutions to complex data problems. She has research experience in graphical models, social networks, demographics, and computer vision. She also served as a visiting student at the University of Oxford, where she studied mathematics.

Yukun Yang, Ph.D. Student

Yukun Yang, a doctoral student in Interdisciplinary Design and Media at Northeastern University (MA) earned his master's degree in information science from the University of North Carolina-Chapel Hill, where he is currently an affiliate at the Center for Information, Technology, and Public Life (CITAP). His recent work was published in Scientific Reports and JAMA Network Open, and focuses on the effect of collective communication practices on social media and how these acts uphold or challenge the dominant narratives, power hierarchies, and hegemonic ideologies.

ORC ID: 0000-0002-6284-5880

Executive Summary

Eyewitness identification plays a critical role in criminal investigations and prosecutions across the United States. However, for decades, its reliability has been questioned due to its significant contribution to wrongful convictions. Specifically, mistaken eyewitness identification has been identified as the leading factor in wrongful convictions, particularly in violent crimes such as sexual assault and robbery. The wrongful conviction and subsequent exoneration of Ronald Cotton in the 1990s epitomized the shortcomings of traditional eyewitness identification practices, emphasizing the need for scientifically- informed practices.

Since the first major set of law enforcement guidelines on eyewitness identification was issued by the National Institute of Justice (NIJ) in 1999, extensive research has advanced the field, with multiple organizations—including the National Research Council (NRC) of the National Academies, the International Association of Chiefs of Police (IACP), and the American Psychology-Law Society (APLS)—issuing updated recommendations. Nevertheless, the adoption of these scientifically supported best practices within law enforcement agencies has been inconsistent, often hindered by a substantial gap between scientific research and field practice.

The Science-Practice Gap

Despite decades of research, a time lag persists between scientific advancements and their translation into law enforcement policies and courtroom practices. Several key factors contribute to this science-practice gap:

- **General Research-Practice Lag:** Researchers have demonstrated that it can take 17 years or longer for scientific evidence to be translated into widespread practice.
- **Divergent Goals and Assumptions:** Scientists, practitioners, policymakers, and the legal community often operate with differing priorities, leading to inconsistent understandings hindering implementation of standard practices in the field.

- **Limited Dissemination:** Scientific findings often remain confined to academic publications, which are not easily accessible to many law enforcement and legal professionals.

- **Complex Real-World Application:** Controlled laboratory research does not always account for the complexity and unpredictability of real-world investigative contexts, and the limited exposure to law enforcement practices may lead scholars to misinterpret how practices are actually carried out in the field.

- **Court decisions:** Additionally, judicial precedents, such as *Neil v. Biggers* (1972) and *Manson v. Braithwaite* (1977), continue to govern legal standards for eyewitness evidence despite being widely regarded as outdated and inconsistent with contemporary scientific evidence, and more recent decisions are also of questionable merit.

Study Purpose and Design

In response to the continued challenges in aligning field practices with scientific evidence, the National Policing Institute (NPI) conducted a nationwide survey of law enforcement agencies, funded by the NIJ. This study serves as a follow-up to the Police Executive Research Forum (PERF) 2013 survey, with the goal of:

- Describing current policies and practices for photo arrays (lineups) and showups (field identifications).

- Assessing changes in practices over the past 10+ years.

- Evaluating law enforcement awareness of recent research findings on eyewitness identification, specifically related to presentation methods and witness confidence.

- Examining the degree of alignment between agency practices and recommended best practices from scientific and professional organizations.

In addition to the national survey, the broader NIJ-funded research project included a field study in two U.S. jurisdictions and a laboratory experiment conducted in the U.K., both examining the relationship between witness confidence and identification accuracy in showup procedures (see Amendola, et al., 2025).

Key Findings

Survey Participation

- A total of 187 law enforcement agencies responded to the survey, yielding a 22.3% response rate, considered more than acceptable by today's standards.

- While lower than the PERF 2013 response rate, the achieved rate exceeds the minimum thresholds for reliable population estimates in large-sample surveys, especially considering declining survey rates exacerbated by the COVID-19 pandemic.

Eyewitness Identification Procedures

- **Photo Arrays:** 86% of responding agencies conduct photo arrays, with higher adoption among larger agencies and those outside the Southern region.

- **Live Showups:** 64% of agencies reported conducting live showups, with significant variation across agency types (showups are significantly more common in local police departments than in state police departments or county sheriffs' offices/departments). Similar to photo arrays, showups are not as common in the Southern region of the U.S.

Collection of Confidence Statements

- 84% reported collecting confidence statements for photo arrays, consistent with previous findings.

- 72.3% of agencies reported collecting witness confidence statements after showups, a decrease from 85% in the 2013 PERF survey.

- Notably, over half of the responding agencies were unaware of recent scientific evidence emphasizing the strong predictive relationship between initial witness confidence and identification accuracy.

- Local police and sheriffs were more likely to be aware of this research (50% +) as compared to state agencies (27%).

- Small agencies were significantly more likely (54%) to be aware of this research than mid- to large-sized agencies (< 30%).

Photo Array Presentation Methods

- There are two types of common photo array presentation methods: ***sequential*** (showing photos one at a time introduced in the 1990s) and ***simultaneous*** (the traditional approach of showing a group of photos at the same time).

- In the current survey, we found that 61% of agencies reported using the sequential presentation method, 24% the simultaneous method, and 13% allowed the use of either method.

- While considerable research in the 1990s and early 2000s suggested an advantage to using *sequential presentation* methods, research over the past decade has either demonstrated no definitive advantage for sequential presentation, or in a number of cases, a slight advantage for the simultaneous method.[1]

- It appears that most law enforcement agencies (LEAs) are not aware of the most recent research, despite the 2014 report of the National Research Council (of the National Academy of Sciences) that did not promote one method over the other and called for additional research. Moreover, the most recent (2016) model policy and guidance from the International Association of Chiefs of Police (IACP, 2020) recognizes "that the sequential and simultaneous presentation methods are both valid methods of conducting an identification procedure and do not recommend one over the other." p. 1.

Policy Development and Adoption

Encouraging progress was observed in the adoption of formal policies:

- 92% of agencies now have written policies governing both showups and photo arrays, a substantial improvement from the 2013 PERF survey, where fewer than half of agencies reported having such policies.

- Nevertheless, gaps persist in areas such as:

    - Standardized instructions provided to witnesses.
    - Documentation of identification procedures.
    - Consistency in how confidence statements are collected and documented.

Barriers to Evidence-Based Practice

The new study identified several enduring obstacles hindering the alignment of agency practices with scientific recommendations:

- **Lack of a Single Authoritative Source of Information:** Unlike in 1999 when NIJ issued the most authoritative guidance on eyewitness practices based on a large task force

[1] This scientific debate hinges on a more refined method deriving from signal detection theory (SDT) that takes into consideration the likelihood of individuals to choose a person from a photo array (known as response bias). SDT accounts for and helps distinguish between response bias from memory accuracy (also known as discriminability). Research has demonstrated that response bias can be minimized through better instructions to witnesses and improved law enforcement.

of scholars, police leaders, and others, today there are numerous sources of information with no single source considered superior to others. Some agencies purchase policies from organizations assuming they reflect the most recent and best evidence (even if they do not), others rely on professional organizations like the IACP or CALEA, and some may also rely on scientific organizations, like the National Academy of Sciences (despite their report being issued over a decade ago), or knowledge of some research articles and findings, but without knowledge of how to judge the most dependable or accurate.

- **Assumptions about Showups being *Too Suggestive:*** Many scholars and advocacy groups have concluded, based on limited evidence, that show-up procedures are too suggestive and therefore should not be used. However, as many respected sources have noted, procedures to reduce suggestiveness across all eyewitness identification practices can/should be adopted. Moreover, those asserting showup suggestibility may not have real-world knowledge about who conducts field showups and how the showups come about. We argue that photo arrays are more likely to be suggestive as they are conducted on official law enforcement property, under the direction of a case detective, after probable cause for an arrest has been established, and where a prosecutor is present in most cases.

- **Inconsistent or Conflicting State Policies:** While some states have established model policies, others lack formal guidance, leading to variability in practice. Moreover, some of the states adopted policies either prematurely or have not adapted those guidelines of policies based on more recent scientific evidence, including the 2014 NRC report.

- **Complexity of Scientific Evidence:** Rapidly evolving and sometimes contradictory research findings make it difficult for agencies (and state authorities) to discern which practices are most supported by the most recent science.

- **Limited Dissemination Channels:** Scientific findings are often not communicated effectively to law enforcement agencies through relevant channels or standardized training.

- **Judicial Influence:** Legal standards, even those regarded as outdated or inconsistent with science, shape agency practices, contributing to resistance or confusion regarding best practices.

Conclusions and Recommendations

The findings illustrate both progress and persistent challenges in the field of eyewitness identification:

- Significant improvements have been made in the adoption of formal policies and procedures over the past decade.

- Despite these advancements, widespread gaps remain in the awareness and implementation of contemporary scientific findings, particularly regarding presentation methods and the importance of immediate confidence statements.

- A substantial science-practice gap continues to undermine efforts to ensure accurate and reliable eyewitness identifications.

- It appears unclear which set of guidelines or recommended practices are the most authoritative especially because NIJ has not attempted to update their 1999 guidance using a similarly comprehensive approach as they used in the late 1990s leading up to that key set of recommended practices. While some of the original guidance continues to be relevant and accurate today, the age of the guidelines has rendered them less authoritative over time.

Broad Recommendations include:

1. **Enhanced Dissemination of Scientific Findings:** Develop targeted, accessible dissemination strategies to communicate research findings to law enforcement, policymakers, and the courts.

2. **Ongoing Training:** Expand training opportunities for law enforcement personnel to incorporate the latest research and best practices in eyewitness identification.

3. **Collaboration:** Engage in partnerships between scientific experts, law enforcement leaders, and legal professionals (such as prosecutors and defense attorneys) to support evidence-based policy development and practice.

4. **Judicial Reform:** Encourage the legal system to reexamine outdated standards for evaluating eyewitness evidence in light of contemporary scientific knowledge using the most authoritative and recent guidance.

Final Remarks

This research underscores the critical need for continued efforts to bridge the gap between science and practice in eyewitness identification. By fostering collaboration, improving dissemination of research findings, and improving training, the criminal justice system can improve the accuracy and reliability of eyewitness evidence, thereby enhancing justice.

A National Survey of Eyewitness Identification Policies and Practices in Law Enforcement Agencies 2023

Eyewitness identification in law enforcement investigations has been an area of widespread concern at least since the 1950s, and even more so since the first DNA exoneration in 1989. In the mid-twentieth century, one legal scholar noted that it had been over a century since a court cautioned about using eyewitness testimony alone without other tangible proof, indicating that most members of the bar have long known of its fallibility (Lasker, 1954). Moreover, in the literature on wrongful convictions, mistaken eyewitness identification is often cited as the most common contributor (see e.g., Connors et al., 1996; G. Wells et al., 1998), especially in sexual assault and robbery cases (see e.g., Garrett, 2011; Gross & Shaffer, 2012; Innocence Project of New Orleans, n.d.). Furthermore, exonerations from eyewitness identification have raised questions about the reliability and efficacy of their use.

Nevertheless, many practitioners have argued that eyewitness testimony is essential in criminal investigations, and scholars have supported that claim. For example, in an interview with leading scholar G. Wells by the American Psychological Association (APA), Wells asserted that "The legal system wouldn't be able to function without eyewitness identification," adding that "we have to find ways to make it better" (Weir, 2016, p. 40). In 1978, G. Wells coined the terms "system" and "estimator" variables to distinguish those important variables under the control of the criminal justice system more broadly (system variables) and those variables that relate more to the context of the crime itself (estimator variables) over which criminal justice system actors have no control. Undeniably, G. Wells and other scholars (see e.g., Wise et al., 2014) have found that misidentifications are often based on problems with the administration of identification procedures or problems that can impact eyewitness memories, such as viewing distance or level of stress, among others.

To be sure, McNabb et al. (2017) asserted that "the reliability of both human memory and the procedures traditionally used by law enforcement for making identifications have been called into question in recent decades" (p. 3). While numerous cases exemplify some of the problems, one specific case from 1984 demonstrates the fallibility of eyewitness memory and how some traditional law enforcement practices may lead to wrongful convictions. The case in question began after a knife-wielding stranger raped then-college student Jennifer Thompson. During the attack, she committed herself to observing and recalling as many details about her attacker as she could. The ensuing investigation, identification procedures used (including a composite sketch and a lineup procedure), assumptions of the police investigators, coupled with both her identification and portions of her statement led to the wrongful conviction of Ronald Cotton (Thompson-Cannino et al., 2010) and his sentence of life in prison plus 54 years (PBS, 2018). A range of problems surfaced after the conviction of Cotton, including the conclusive DNA identification of Bobby Poole (a known sex offender) as the perpetrator, leading to Cotton's exoneration and pardon by North Carolina's governor in mid-1995 after serving 10.5 years of his sentence.

Among the problems in the case were the conditions under which Ms. Thompson viewed her attacker, the incomplete medical examination she received at the first hospital, the flawed procedures the police used in the development of a composite sketch and the live lineup, the limited corroborating evidence, and Thompson's own highly confident identification at the trial. Yet, in her first viewing of photographs, Thompson took several minutes going back and forth between two possible suspects (Weir, 2016), and when she finally chose Cotton, she was less than certain. Yet, she became more confident over time, leading up to the trial, at a time in history when courtroom identifications were both common and widely accepted. These factors

might have alerted police that this was hardly conclusive evidence, but scientific knowledge about eyewitness identification was at a much earlier stage than it is today.

For example, in the 1980s, there were no widely accepted law enforcement guidelines, although model rules on eyewitness identification had been published by the Police Foundation in the 1970s (Arizona State University, College of Law, 1974).[2] One purpose of those model rules was to "... promote practices which will safeguard innocent persons while insuring that those who are guilty are arrested and convicted" (p. 3), a goal that is very much at the core of scholarly research over the last half century. Indeed, balancing the need to bring the guilty to justice and to avoid convicting the innocent takes precision in approaches and methods, as well as accuracy in our tests of reliability and accuracy.

Undoubtedly, much has been learned through cases such as Thompson's (and many others), ongoing scientific research, and authoritative guidance provided by a number of organizations, perhaps most notably the National Institute of Justice (NIJ) *Eyewitness Evidence: A Guide for Law Enforcement* (1999) (hereafter referred to as the NIJ Guidelines) and the National Research Council (NRC) report, *Identifying the Culprit: Assessing Eyewitness Identification* (2014) (hereafter referred to as the NRC Report).[3] For example, scientific consensus now includes the fact that courtroom identifications are not useful or reliable (see e.g., Garrett, 2011; Wixted et al., 2015), faster identifications (response time/latency) are more likely to be accurate (Dodson & Dobolyi, 2016; Quigley-McBride & G. Wells, 2023; Seale-Carlisle et

[2] In 1983, The Law Reform Commission of Canada also issued their own Guidelines (see N. Brooks, 1983).
[3] The NRC Report was generated by the National Research Council's Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts, Committee on Science, Technology, and Law, Policy and Global Affairs' Committee on Law and Justice Division in the Behavioral and Social Sciences and Education, for the National Research Council of the National Academy of Sciences.

al., 2019), and that immediate confidence statements are predictive of accuracy,[4, 5] whereas later ones are not (see e.g., Wixted & G. Wells, 2017).

Guidelines for Eyewitness Identification Procedures

A few years after the exoneration of Mr. Cotton, NIJ convened a working group (made up of experts and practitioners) to develop, evaluate, and recommend a set of procedures to reduce the likelihood of false identifications, many of which are still followed today, some 25 years later. The working group's 1999 report, *Eyewitness Evidence: A Guide for Law Enforcement*, offered recommendations to LEAs, including the following that still appear to be among today's best practices[6] that are recommended and/or supported by many scientists, professional organizations, and/or the 2014 NRC Report. These include, but are not limited to,[7] the following.

1. Develop policies and procedures for eyewitness identification of suspects.
2. Develop and/or implement standardized instructions to witnesses.
3. Record witness statements.
4. Caution the witness that the person ... may or may not be the perpetrator.
5. Obtain and document a statement of confidence/certainty.
6. Train all law enforcement officers in eyewitness identification.[8]

---

[4] Sauer et al. (2019) cautioned that the confidence-accuracy relationship at first viewing is justifiable at the aggregate level, it may not be when attempting to evaluate any single identification, yet this is true for all probability statistics.
[5] Wixted & G. Wells (2017) noted that the relationship relies on "pristine testing conditions," but Mickes et al. (2017) note that witness confidence overshadows that concern in terms of the identification being probative of guilt.
[6] The term "best practices" infers recent evidence-based practices, but it is referred to somewhat loosely across fields (see, e.g., Veselý, 2011). For more information about its meaning and the use of the term, see the "Discussion" section of this report.
[7] Recommendations about using "blind," "blinded," or double-blind procedures in administering procedures has also been covered by the 1999 NIJ Guidelines and subsequent recommendations. However, how these practices are administered is often confounded with lineup presentation methods (sequential versus simultaneous presentation). We used a different question than that used by PERF in our survey.
[8] Few details about the specific curriculum for eyewitness identification training have been provided other than the NIJ trainer's manual, see the archive version from the Department of Education.

More than a quarter of a century has passed since NIJ established eyewitness identification guidelines. While many of the NIJ 1999 Guidelines are still relevant today, the narrative surrounding eyewitness identification has changed significantly over the last 25 years, due to increased scholarly activity and new evidence related to eyewitness identification reliability and accuracy. More recent research findings across several areas of eyewitness identifications have resulted in changes and/or additions to the generally accepted knowledge among scholars and other experts in the field, and there are considerable inconsistencies.

New Sources of Guidance on Eyewitness Identification Practices

Although the NIJ Guidelines were the primary authoritative source of guidance over 25 years ago, many highly regarded organizations and scholars have issued new guidance since then. In 2013, the Police Executive Research Forum (PERF) published *A National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies*, which was funded by NIJ and documented policies and practices adopted by law enforcement agencies. Shortly thereafter, the NRC Report was published. It contained a detailed summary of the relevant scientific evidence from laboratory and field studies in eyewitness identification and memory science and issued new recommendations. Importantly, its review of the scientific evidence on eyewitness identification covered the science that emerged over the past 30 years (including research conducted in the 15 years prior to the 1999 NIJ Guide and the 15 years subsequent).

That report and set of recommendations represented current science and practice through 2013 (and part of 2014). Perhaps at least partly in response to those recommendations, the International Association of Chiefs of Police (IACP) updated its guidance, "Model Policy: Eyewitness Identification," in 2016,[9] and in 2017, the U.S. Department of Justice issued a memo,

[9] IACP Policy Center, 2016.

"Eyewitness Identification: Procedures for Conducting Photo Arrays,[10] reflecting some of the best practices supported by research. For those agencies with national accreditation through the Commission on Accreditation of Law Enforcement Agencies (CALEA), their standards related to lineups and showups have also been updated during this period. More recently, a subcommittee of the American Psychology-Law Society (APLS or APA's Division 41) published an updated review paper (see G. Wells et al., 2020) delineating new guidelines based on feedback from some of their members and the legal community. The extent to which it has been endorsed more broadly is not evident.[11] However, while the American Bar Association (ABA) issued its own best practices 20 years ago, its recent book on eyewitness reform does not specify best practices per se [12] (see Collins, 2024).

Additionally, according to Albright & Garrett (2022), at least 26 states have established their own guidelines, recommendations, and/or 'model' policies and procedures for eyewitness identification, dating from 2001[13] through at least 2022. Moreover, some courts have issued decisions that, in some cases, are inconsistent with scientific evidence. For a thorough discussion of state statutes, state and federal court rulings, and model policies, see Albright & Garrett (2022).

Field Adoption of NIJ's 1999 Eyewitness Identification Guidelines

To examine the extent to which the NIJ Guidelines had been adopted by law enforcement agencies nationwide, and to explore practices being employed in the field, the PERF conducted a nationwide survey of eyewitness identification (ID) practices in the field well over a decade ago

[10] Known as the Yates memo on eyewitness identification (2017).
[11] As per the authors, they "do not represent the position of the APA or any of its other Divisions or subunits." (p. 3)
[12] The publisher touts this as a comprehensive guide to make the process fairer and more accurate. Free to ABA members only.
[13] New Jersey guidelines were established in 2001.

(PERF, 2013). Funded by a grant from NIJ, that survey and its results have undoubtedly proved useful for law enforcement agencies (LEAs), scientists, and policymakers alike.

Among the key findings of the PERF report were the following:

- Most LEAs had not fully implemented all the NIJ 1999 recommendations.
- Some of the 1999 NIJ Guidelines had been employed in many agencies, while others had been implemented in less than half of the surveyed agencies.
- About two-thirds of LEAs had no written policies for photo arrays, and over 75% had no written policies for showups.

Recent Issues Related to Eyewitness Identification Practices

Some specific research areas on eyewitness identification reliability, methodology, and accuracy have been more closely examined by scholars since or just before the results of the PERF survey. Among the four most explored and debated issues over the ensuing period (and in some cases, previous period) are (a) questions surrounding the most appropriate presentation method for photo arrays (i.e., simultaneous *versus* sequential presentation methods); (b) the ability of witnesses to differentiate between guilty and innocent suspects (or discriminability); (c) a new analytical approach (receiver operating characteristic) to assess discriminability by also considering the extent to which individuals are more or less inclined to pick someone from an identification procedure (called response bias): and (d) the relationship between witness confidence and accuracy, including a new method called confidence-accuracy calibration. Some of that research has upended previous findings, suggesting the need to consider newer methods, analyses, and research.[14]

[14] For a discussion of these issues from a prosecutor's perspective, see essays by Riley (2018, 2019), former Assistant U.S. Attorney of the Washington, DC U.S. Attorney's Office.

Problem Statement

Despite decades of scientific research on eyewitness identification, field-based policies and practices, and judicial decisions have not always followed or kept pace with scientific findings. Several reasons exist for the lack of consistency between scientific findings, policy adoption, and practice improvements. Undoubtedly, a sizable and overarching science-practice gap has been noted by many scholars across fields (see e.g., Norman, 2010; Morris et al., 2011). In some cases, this is due to counterintuitive findings or anecdotal evidence. However, many of the common reasons for the disconnect are the average time lag between research and practice, differing goals and assumptions by scientists, practitioners, and the courts, and the lack of appropriate or sufficient dissemination of key scientific findings to those responsible for implementing policies and practices at various levels. The evolving nature of science and inconsistent or unreplicated scientific findings can lead to considerable confusion over which research to rely upon when establishing or updating policies and practices. Some of these inconsistencies are rooted in differing methods, measures, and levels of scientific rigor.

The Science-Practice Lag

A key reason for the science-practice gap is the rarely considered time lag between research and practice. It has been nearly a century since Stoddard (1935) opined that practice lags behind research. While the length of the *science-practice* time lag has not been well-established in psychology or criminal justice, in the field of medicine, several scientists have suggested that the time lag averages about 17 years (see e.g., Morris et al., 2011; Rubin, 2023) and one scientist in criminal justice, argued that it may be at least that long, if not longer in criminology (see Blomberg, 2021).

Indeed, many reports, policies, and training outlines often refer to standard practices

based on scientific findings that have since been disputed on the basis of new evidence. For example, in examining more recent published articles about adopting best practices in eyewitness identification, some considered only the 1999 NIJ Guidelines as the sole source for best practices (see Bertrand et al., 2018), despite that guidance being issued 25 years ago. Yet, research across disciplines is rapidly evolving, and in eyewitness identification, the evolution has been particularly evident over the last 20 years. While scientific research on eyewitness identification has moved apace, it has moved increasingly rapidly in recent years. Expecting rapidly advancing science to be adopted more quickly in the field may be unreasonable given evidence about the slow pace of diffusion (see e.g., Green et al., 2009).

Differing Goals and Assumptions Among Experts

Green and colleagues (2014) argued that "the prevailing disappointment with the flow of scientific information into practice has much to do with misguided expectations ... that the truths discovered by science ... should automatically influence behavior" (p. 4). This belief is likely rooted in the faulty assumption that scientists and practitioners have the same goals and understandings, and rely upon each other to do their work, or, for that matter, interact in meaningful ways. Undoubtedly, there are those who do not believe science should inform either field-based practice or the courts, instead relying on personal, lived experiences, judicial precedents, or personal (often unsubstantiated) beliefs. However, even among those who do support the use of science in policymaking, practice (i.e., evidence-based practices), and/or the courts, there are undeniably differing assumptions, beliefs, and goals among various researchers (from academia vs. other settings), policymakers (at the local, state, and federal levels), practitioners (police chiefs, sheriffs, and investigators), and the courts (including attorneys adjudicating cases or representing defendants, judges, and juries).

A key consideration in understanding some of the underlying differences, is that purely controlled laboratory research often cannot neatly fit the complexities of the field, or as Norman (2010) eloquently articulated " ... the real world is complex and messy, with uncontrolled variables, sometimes behaving in ways that contradict the neat, tidy, logical assumptions of the scientist" (p. 11). When conducting field-based research, some researchers try to balance the rigor of lab studies with the practical realities of the field. Still, their work may be rejected by laboratory scientists as not being as rigorous or lacking the specific controlled laboratory environment.

Similarly, many practitioners with lived experiences and intuitions may rely more heavily on their knowledge gained from those than on complex, scientific studies that have not been clearly or concisely translated for their use. Accordingly, some concerns about the external validity of the research are raised when translating a lab study into the field. Also, attorneys and judges may rely more heavily on legal precedent and legal arguments, and jurors may depend on their own life experiences and the arguments presented in court when interpreting scientific arguments, due to their limited exposure to many research findings. And so, perhaps not surprisingly, Green et al. (2009) further argued that "the blame for gaps between science and practice points variously at tradition-bound practitioners, who insist on practicing their way and believe they know their patients or populations best, and at the smugness of scientists believing that if they publish it, practitioners and the public will use it." (p. 154)

Complex Relationship between Science and Law

According to a neuroscientist who is a leading forensic expert and authority in eyewitness identification, there is a "complex and sometimes contentious relationship between science and law" (Albright, 2023). Decisions and associated opinions by courts often serve as a source of law

enforcement policies and procedures, although they do not always reflect scientific evidence.[15] Most notably, the Supreme Court standards established in the 1970s under both Neil *v.* Biggers, 409 U.S. 188 (1972) and Manson *v.* Braithwaite 432 U.S. 98 (1977)[16] remain the constitutional standard after decades of contradictory scientific evidence, leading some researchers to repudiate these precedents.

For example, McNabb et al. (2017) asserted that "the U.S. Supreme Court and a significant majority of state supreme courts apply a test for eyewitness reliability that has no foundation in empirically validated science and predates 40 years of biological science and social science research" (pgs. 16-17). Furthermore, they noted that few jurisdictions have updated their legal frameworks related to eyewitness identification despite newer scientific evidence, and even those that have done so have changed the admissibility of eyewitness evidence in court or mandated evidence-gathering procedures for LEAs. This is not at all an isolated view, as the NRC (2014) noted that "not all of the Manson factors are diagnostic of reliability" (p. 6).

While many LEAs rely on courts and case law to provide the best evidence, it is important to note that attorneys adjudicating cases may also rely on individual expert testimony regarding current scientific findings. Undoubtedly, not all experts called upon in these cases have necessarily kept abreast of the current scientific findings,[17] and even if they have largely done so, they are typically unlikely to present conflicting evidence, depending on their role in the case at hand. The adversarial nature of our justice system means that there are often competing experts, leaving juries and judges to sort out the most compelling arguments, not necessarily the most

[15] This phenomenon is not isolated to legal issues in eyewitness identification. For example, in a recent editorial in *Scientific American* on Shaken Baby Syndrome, the authors noted that "the U.S. legal system often fails to recognize advances in scientific knowledge" (Kukucka & Faigman, 2024, subtitle, p. 1).
[16] For a discussion on the U.S. Supreme Court Ruling in *Manson v. Braithwaite*, 432 U.S. 98 (1977) see McNabb et al., 2017.
[17] This clearly is not an easy task even for seasoned scientists or lawyers, as the body of research is extensive and dynamic.

accurate or relevant scientific findings. While the adversarial process does provide some safeguards, it does not guarantee that the science underlying each argument presented is based upon fully accurate, recent, rigorous, or high-quality research. Importantly, the nature of scientific discourse is such that experts do not always agree, and scientific findings regularly evolve (see e.g., Albright & Garrett, 2022; Lombrozo, 2017), leaving justice-related decisions sometimes to be rendered on incomplete or outdated scientific findings, *fait accompli*.

The Lack of Sufficient or Appropriate Dissemination of Scientific Findings

The lack of sufficient and appropriate dissemination, especially to those responsible for implementing evidence-based practices, may cause the evidence gained from scientific research to remain solely with scientists and their colleagues. Research scientists, particularly those in academia, are primarily responsible for research and publication in scientific journals. Accordingly, these scholars are typically not incentivized to translate or present their findings to the practitioner or policy communities (see e.g., Rouzer et al., 2023). At the same time, there is considerable debate about the extent to which scientists should advocate for policies. As such, some scholars may see their roles as scientists as incompatible with advocacy and therefore, may avoid providing policy or practice guidance to the field (for a discussion, see e.g., Hinojosa & Nesterak, 2021).

Moreover, articles in research journals often contain scientific jargon, complex theory, and/or complicated methods and statistical analyses, making them difficult for policy and practitioner audiences to interpret. In addition, some published reports and guidelines are lengthy. For example, the NRC's comprehensive report is 154 pages, and its content is very technical. Such a lengthy, detailed publication with complex nuances and analytical procedures may have been difficult for policymakers to interpret and understand. It thus may not have been

entirely relied upon as a source of policy adoption or change at the state or local levels, even if its availability was well-known in the field. It is perhaps less likely that police personnel would have run across the APLS guidelines, since they were published in a psychology journal, farther afield from typical policing or criminal justice publications.

Importantly, dissemination efforts by researchers, their organizations, institutions, or funders are not typically visible, other than through publications in academic journals. Hence, the extent to which comprehensive, guided dissemination efforts have been implemented is often unknown. Nevertheless, there are many avenues and means for informing policy and practitioner audiences, such as presentations at professional conferences (not just research conferences), electronic distribution of research findings and/or reports, mailouts of printed materials, public media like news reports or publication in mainstream magazines and newspapers, research institution special or annual reports, social media, blogs, webinars, and trade publications (for police, sheriffs, prosecutors, defense attorneys, judges, etc.).

Adoption of Recent Evidence and Recommendations

To date, there is a lack of available information about whether the NRC and APLS reports have informed state guidance, laws, or training academies, or were disseminated or summarized through more direct channels, such as through policing conferences or in publications and/or magazines that police leaders and officers more readily follow. This gap suggests that LEAs and their personnel may be unaware of the newest evidence.[18] Because these two reports were published quite recently (i.e., considering the typical science-practice lag), it is unlikely they have been widely disseminated or adopted. Nevertheless, Albright and Garrett

[18] With thousands of scientific journals, and certainly dozens focused on law, law enforcement and criminal justice, it is unclear which, if any, are reviewed by law enforcement agencies. Moreover, psychology journals are even less likely to be readily available to, or shared with, those in law enforcement.

(2022) asserted that the NRC Report "has informed a remarkable shift in practices" (p. 577). Yet, there is little evidence to support that claim in terms of field-based practice, other than an isolated study by McNabb and colleagues (2017) in which the authors examined a five-state region of the U.S. (four midwestern states and one southern state). Even so, these researchers included some recommendations that had also been part of the original NIJ Guidelines, so determining whether policy adoption came from NRC or NIJ's recommendation is unknown.

Additional Barriers to LEA Policy & Practice Guideline Adoption

Despite extensive research studies on eyewitness identification, including systematic reviews and meta-analyses, not all studies or publications carry equal weight due to their quality, methodology, and scientific rigor. Likewise, not all scientific journals in which studies are published are of the same quality or impact due to their scientific reputations, impact factors,[19] potential publication bias, or biases of editorial board members,[20] among other considerations.

Due to the lack of scientific consensus on some issues, various methods, and inconsistent findings across studies, it is almost insurmountable for LEAs to sort through and identify which recommended policies and procedures are most accurate or current. Plus, perhaps most importantly, the guidance provided to LEAs often depends on the source of the guidance, and inconsistencies abound.

[19] There are a number of rating systems to identify the "impact factors" of various journals, and some journals are widely believed to be top-tier journals in the field in the social sciences.

[20] There are several forms of editorial bias discussed in the literature. One that is seldom mentioned is the long tenure of editorial board members, i.e., no term limits for editorial board members which may limit diversity of thought and/or perspective, sometimes biasing what gets published when it conflicts with their members' own research findings or views. Moreover, some editorial boards promote the use of regular authors as editorial board members of the same journal, despite the likely conflict or bias; something that seems to be considered acceptable anyhow.

Current Studies

With all the aforementioned factors in mind, researchers at the National Policing Institute[21] (NPI) and their partners submitted a proposal to NIJ in 2019 in response to a competitive solicitation under the "Administration of Justice" category. The proposal, which was selected for a grant award, consisted of multiple studies to examine policies and practices and to advance science on the relationship between confidence and accuracy in show-up procedures (also known as field identifications).

The proposed studies included a new nationwide agency policy and practice survey as a follow-up to the PERF study published over a decade ago, and to address new questions related to more recent research. The other two studies that were part of the same NIJ-funded grant project addressed the relationship between confidence and accuracy in showup procedures, with a field study in two U.S. jurisdictions and a laboratory experiment conducted in the U.K.[22]

The National Policing Institute Eyewitness Policies and Practices Survey

This survey addresses eyewitness policies and procedures adopted since the PERF and NRC reports and recommendations, both published over a decade ago. The survey was seen as critical since considerable research had amassed since the 2013 PERF survey on field-based practices, much of which was reviewed in the 2014 NRC Report and served as the basis for their recommendations, and because there had been limited evidence about field-based changes in eyewitness practices.

While undoubtedly some state and other policy development and guidance have been

[21] Formerly known as the Police Foundation
[22] Descriptions of the combined laboratory and field studies and their results are provided as a separate report (Amendola, Mickes et al., 2025).

updated since 2014, little is known about a change in *actual* field-based practice, which often deviates from best practice recommendations. One exception is a survey by McNabb and colleagues (2017), in which they examined the translation of eyewitness guidance to the field across five states.[23] A few of their findings were similar to or consistent with those found by PERF (2013), including the following: (a) inconsistency in instructions provided to witnesses; (b) lack of consistency of eyewitness research methods across agencies; and (c) few challenges to face when implementing policy changes.

Purpose, Goals, and Objectives

The goal of this NPI report is to summarize LEA policies and practices associated with photo arrays and showup procedures and assess the extent to which current policies and practices reflect the current state of the science and/or recent guidance. The specific goals are: a) comprehensively describe current policies and practices associated with photographic lineups (arrays) and showups (field identifications), (b) compare results from our survey to the prior results from the PERF survey (where possible), (c) examine the responses to questions with more refined response options, (d) identify the sources of agency policies and practices, (e) examine the extent to which policies and practices are consistent across agencies, and (f) assess LEAs' knowledge of some recent scientific findings over the past decade regarding presentation methods and the relationship between confidence and accuracy.[24]

Research Questions

The investigators sought to address the following research questions:

[23] Arkansas, Iowa, Kansas, Missouri, and Nebraska (the former in the South, and the latter four in the Midwest)
[24] Current scientific findings refer to those during the past decade and represent general scientific consensus among experts in memory and eyewitness identification. Consensus may sometimes be unclear, especially when there are inconsistent findings across studies; where the current science is unclear, we provide explanations throughout.

1. To what extent are LEA agency policies and practices regarding eyewitness identification consistent overall and by agency size and type?
2. Have current LEA practices and procedures for photo arrays and showups changed overall and by agency size and type over the past decade?
3. Which sources do LEAs rely on to inform eyewitness policies and practices?
4. To what extent are agencies knowledgeable about recent scientific findings about eyewitness presentation methods (sequential versus simultaneous) or the relationship between immediate confidence and accuracy?
5. To what extent are LEAs' policies and practices consistent with recommended practices?

Overview

The issue of the utility of research in informing practice underlies all three research studies conducted as part of this NIJ project: This report, which describes the national survey conducted with law enforcement agencies nationwide; a study consisting of a laboratory experiment conducted in the U.K.; and a field-based evaluation in two U.S. jurisdictions designed to address whether eyewitness confidence and identification latency are predictive of accuracy of identifications in showup procedures.

While the lab study was conducted using an experimental paradigm, the inability to manipulate certain real-world conditions ethically and practically meant that the field study relied on previously collected (archival) data from two law enforcement jurisdictions where witness confidence statements in showup identifications were documented.

The field studies also relied upon law enforcement practitioners (both an experienced police investigator and prosecutor) to evaluate the strength of the corroborating evidence in cases where confidence statements were obtained from witnesses. Using both a lab and field study to examine the same basic hypotheses was seen as optimal to bridge the science and practice divide and better understand field and lab-based limitations in addressing issues in showup practices.

Research Design, Methods, and Analysis

Key personnel at the NPI conducted the Nationwide Survey of Policies and Practices in Eyewitness Identification (hereafter referred to as the NPI Eyewitness Policies and Practices Survey) as a 10-year follow-up to PERF's National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies (PERF, 2013).

Survey Development, Content, and Format

We modeled the survey largely after the decade-old PERF survey. To construct the new survey, we consulted with additional experts to examine prior question wording (to determine if items or response options should be changed to reflect newer information or to get more refined responses) and pose additional questions in response to newer research findings and unresolved issues.

This report contains the results from the new NPI nationwide survey.[25] It differed from the prior survey in several unique ways. Firstly, unlike PERF, our survey was web-based,[26] while theirs was distributed via U.S. mail[27] and included follow-up phone calls. We used the Qualtrics® survey platform to distribute the web-based survey, and for those agencies for which we did not have an email address, we mailed a letter containing a link to complete the survey online.[28] Secondly, the NPI survey did not include any phone interviews. These differences were primarily due to survey length, cost, and technological advances in survey administration,

---

[25] This report served as a 10-year follow-up to PERF's survey (2013); the new data was collected almost exactly 10 years after the PERF survey. However, delays associated with Covid-19, and the fact that there were three different studies covered by this grant, led to an approximate two year delay in publication of this report.

[26] Using Qualtrics® survey software; 75 agencies sampled from the NDLEA did not have associated email addresses, so we contacted them by U.S. Mail (USPS) and provided a short version URL link to complete the survey electronically. In addition, after we sent the electronic invitation to the sampled groups, some of the emails bounced. There were 47 invitations for which the email bounced, and we sent mail invitations via USPS to those agencies as well. What we do not know, however, is how many of these mailed invitations reached the agencies.

[27] These paper surveys were machine readable (using a form called Teleform). These were distributed over three waves during a period of just over two months, followed by five sets of reminders over about three months.

[28] As evidenced by a bounced back email.

distribution, and analysis over the past decade.

Our survey addressed only two eyewitness procedures: live showups[29] (also known as field identifications) and photo arrays (often referred to as photo lineups), as compared to PERF's survey, which also included questions about composite sketches, live lineups, and mugshot searches. We excluded the latter practices to reduce agency burden, because they are utilized less frequently,[30] and because major advocacy organizations are less likely to recommend them. Also, the NRC Report did not address composite sketches or mugshot searches, and the current IACP policy does not address composite sketches and deals very minimally with mugshots.

Since we were interested in the most recent scientific findings, we also compared our policy and practice results to those obtained by PERF over a decade ago, and included recommendations covered by the NRC (2014), the IACP (2016), and the APLS report (G. Wells et al., 2020). Specifically, the NPI survey included questions surrounding the growing body of research on photo arrays and showups from the past decade, emphasizing presentation methods and the initial confidence-accuracy relationship.[31]

Although the survey included similar questions, it was also organized differently from the PERF version. Our survey contained two primary sections addressing the most commonly used eyewitness procedures only—showups and photographic arrays. In PERF's paper, the questions were each asked once, with response options for each of five identification procedures (to also

[29] Referred to by some agencies as field identifications, but this terminology sometimes means something else in other agencies.
[30] PERF found that live lineups were used in 21.4% of agencies, mugshot searches in 28.8% of agencies, and composite sketches in 35.5% of agencies, as compared to photo lineups (94.1%) and showups (61.8%).
[31] Many studies over the past decade have relied on "receiver operating characteristic" (ROC) methodology and analysis (derived from signal detection theory) to compare sequential vs. simultaneous presentation methods. Per the NRC report, ROC is a more comprehensive approach than the traditional *diagnosticity ratio,* and that has advanced the debate about the superior practice. The relationship between initial confidence and accuracy is also now broadly accepted.

include composite sketches, live lineups, and mugshot searches). Moreover, some of our questions were asked differently because of differences in the procedures or to clarify our understanding further.

In both sections of the new survey, the questions included content similar to the previous survey,[32] e.g., current practices, instructions to witnesses, training in identification practices for law enforcement personnel, etc. Importantly, we started with showups, and if an agency indicated that they did not conduct them, they were routed to the section on photo arrays. Finally, we added a few new questions and revised response options based on our new research questions, more recent scientific findings, and/or to provide more response options.

Selection of Sample

We relied upon two sources to determine which agencies to survey: the National Directory of Law Enforcement Administrators (NDLEA, 2020) and the complete database of respondents to the PERF survey (2013) obtained from PERF's research division. Specifically, we used the 2020 NDLEA to identify three distinct population samples (see **Appendix A**, NDLEA Strata Sizes with Selected Sample Count):

1. State law enforcement agencies (e.g., state police, highway patrol, or investigative bureau)
2. Local police departments (e.g., town, city, county)
3. County sheriff's offices/departments

The population of state agencies included all identified state police, highway patrol, and state investigative units. The population of local police departments included city, town, and county police, but not special police like transit or university police. The population of county

[32] We addressed reduction of bias via double-blind, non-blind, and 'blinded' procedures somewhat differently than PERF did.

sheriff's offices or departments (excluding city or town sheriffs) included those agencies that conduct routine law enforcement and/or investigative functions. Consistent with PERF's original sampling strategy, our survey sample was stratified by agency type (local police, county sheriff, state agency), agency size (number of sworn personnel), and geographic region as designated by the census (Midwest, Northeast, South, and West). The Census regions and states within them are presented in map form in **Appendix B**: States with Census Regions.

Once the respondent samples were identified, we used an additional screening question to confirm whether the survey questions applied to the agencies. If the agencies reported not conducting showups or photo arrays, they were routed to the end of the section or survey.

Sampling Methodology and Stratification

The first step of our sampling procedure was to clean and filter the NDLEA database, which consisted of 16,164 agencies in the state and local law enforcement administrators' version. However, this database included some non-policing agencies and other agency types besides sheriffs and local police departments, and there were also some agencies for which multiple administrators were listed. As such, we started by removing duplicates (deduplicating) based on agency name, city, and state, removing non-policing agencies, and filtering to include just local police (town, city, county) and county sheriffs that are engaged in investigative practices.[33] Thus, while there were 16,164 agencies in the NDLEA file, the final database was pared down to 14,190, as shown in **Appendix A**.

The entire survey pool was expected to consist of about 1,400 agencies, which is consistent with the prior PERF survey. However, we split our agency pool into two groups: (1) a nationally representative sample (n = 765) and (2) agencies that had previously responded to the

[33] In addition, our pilot sample was drawn before the sampling procedure, reducing the population slightly (n= 27).

PERF survey (n = 619). Therefore, the sampling strategy for this survey was designed to target a stratified, random sample of approximately 800 representative agencies (stratified, random sample). There were 48 strata representing four regions, each with two agency types, and five agency sizes (plus a stratum that included agencies for which the NDLEA did not include the number of sworn personnel classified as "missing").

The sample was constructed by selecting all agencies in the 17 strata with 30 or fewer agencies, then sampling 20 agencies from each remaining strata (n = 31), weighted by population size. This resulted in a sample of 769 local police and sheriff agencies (see **Appendix A**), all of which were sent a survey. However, because four were undeliverable (returned a response of undeliverable as addressed or similar), our sample denominator for calculating the response rate is based on 765 surveyed agencies (for the second and third sampling frames, i.e., the local police and county sheriffs). We also surveyed all 76 state police, highway patrol, and investigative agencies, of which two were undeliverable, for a total of 74. Accordingly, our final survey pool of ***eligible respondents*** consisted of 839 agencies. When including the other distribution group (the PERF respondents), our survey was distributed to 1,458 agencies.

Survey Distribution

The survey was administered from July 2021 through November 2021 to three groups: (1) the stratified, random sample of local law enforcement agencies (n = 769), (2) the population of state agencies (N = 76), and (3) the respondents to the 2011 PERF survey (N = 619). The former two groups consisted of 845 agencies in total. However, with six undeliverable surveys, our final sampling frame was 839 for groups 1 and 2.

Results from our final survey distribution group—those who had previously responded to the PERF survey—are not included in this report. The purpose of surveying the PERF group was

to track that cohort and compare their responses over time. However, this report focuses solely on the new stratified, random sample of LEAs plus all state agencies.

Since not all agencies in the NDLEA had an associated email address (n = 75) or the emails bounced back to us (n = 6), we used two different approaches (online and paper invitations) to distribute the survey. The paper version was distributed by U.S. mail and contained a survey invitation with a condensed version of the URL link to complete the survey electronically. We do not know how many mailed invitations failed to reach a respondent or agency.

A breakdown of our sample by each stratum is presented in **Tables 1-3**. Shown in **Table 1** are those surveyed by agency type, **Table 2** agency size, and **Table 3** geographic area.

Table 1. *Surveyed agencies by agency type*

Agency Type	Number Distributed
Local police departments	417
County sheriff's departments/offices	348
State police and investigative units	74
Total	**839**

Note. Total represents the denominator for calculating response rates.
[a] Distributed 769 but 4 were undeliverable. [b] Distributed 76 but 2 were undeliverable.

Table 2. *Surveyed agencies by agency size*

# of sworn officers	# Local police & sheriffs	# State Agencies	Total
25 or fewer	159	n/a	159
26-50	159	n/a	159
51199	163	n/a	163
100-499	160	n/a	160
500 +	47	n/a	47
Unknown	77	74	151
Total	**765[a]**	**74[b]**	**839**

Note. Total represents the denominator for calculating response rates.
[a] Distributed 769 but 4 were undeliverable. [b] Distributed 76 but 2 were undeliverable.

Table 3. *Surveyed agencies by geographic region*

Agency Region	# Local police & sheriffs	# State agencies	Total
Midwest	212	17	229
Northeast	180	10	190
South	178	26	204
West	195	21	216
Total	**765**[a]	**74**[b]	**839**

Note. Total represents the denominator for calculating response rates.
[a] Distributed 769 but 4 were undeliverable. [b] Distributed 76 but 2 were undeliverable.

Pilot Test

To pilot the survey, we removed the geographic strata and sampled three agencies from the remaining nine strata (there were no agencies in the County Sheriffs group with missing size information strata), yielding 27 agencies in the pilot sample. Note that this pilot sample was drawn before the updated sample described above due to an early oversight in which we neglected to remove some duplicated agencies from the NDLEA file. The pilot agencies were removed from the eligible population pool before drawing an updated sample.

For five of these agencies, we had no email address, one had an incorrect one, and one was a bounce back. After searching online, we were unable to track down the missing emails. The survey was ultimately sent to 20 agencies via email with one email reminder and one reminder phone call. Three agencies (15%) returned the survey with responses and comments that were considered in finalizing the survey. These three agencies were the Newtown (CT) Department of Police Services, the Fulton County (GA) Police Department, and the Habersham (GA) County Sheriff's Office, all of whom fully cooperated.

Survey Response Weighting

Weighting is a statistical process designed to ensure that survey respondents accurately represent the larger population. It allows the results to better represent the characteristics of the

sample, in this case, agency type, size, and region. By adjusting for underrepresented groups or varying response rates, weighting minimizes bias and increases the reliability and validity of the survey results.

Due to the separate survey frames for local, county, and state agencies, the results presented herein are reported differently for the representative sample than for the state agencies. Weights were assigned to the local police departments (town, city, or county police) and sheriffs' offices or departments (both referred to throughout the report as "local police" and "sheriffs"). State agencies were not assigned weights since that group consists of all the identified agencies selected from the NDLEA rather than a sample. Specifically, the representative sample frequencies were assigned weights to allow for greater generalizability. In contrast, the state agencies' frequencies are *unweighted* as they are based on the entire identified population of state agencies (N = 76). **Appendix C**, Weighting Strategy, discusses how the weights were calculated. Across tables, the frequencies and proportions may not equal the total responses in the sample due to weighting and rounding.

Results and Findings

About Survey Response Rates

The prevailing view among social scientists, at least until the last two decades, was that surveys required at least a 60-80 percent response rate to be reliable, with the greatest concern being response bias. Response bias, or more accurately, non-response bias, occurs when a variable in the survey is correlated with the willingness to respond, meaning respondents and non-respondents differ in systematic ways. However, newer data have upended this standard.

At the same time, it has been well-established that response rates to all types of surveys have dwindled considerably in recent decades. In a study funded by the U.S. Department of

Health and Human Services, the authors reported that across eight major federal surveys, response rates declined during the period from the mid-1990s through 2014, with 75% exhibiting "accelerated rates of decline in the most recent years" (Czajka & Beyler, 2016, p. viii). Moreover, the Covid-19 pandemic led to response rate decreases that were "...larger than in the past" (Ash et al., 2022, p. 3). For example, in a review of six major panel studies in the U.S.[34] between 2019 and 2020, researchers found that for all but one, the average response rate reduction over the period was 29% (Krieger et al., 2023). It is not clear if survey response rates have returned to prior levels since then. Although our proposal to conduct this survey was submitted in 2019, prior to the Covid-19 pandemic, we had no choice but to administer the eyewitness survey during the peak of the pandemic. Accordingly, we expected a lower response rate than anticipated when the proposal was submitted.[35]

The type of survey also influences response rates, e.g., evidence demonstrates response rates tend to be the lowest for web-based surveys, as compared to face-to-face interviews, phone, and mail surveys (Macias et al., 2008; Shih & Fan, 2008). In a recent analysis by Wu et al. (2022), the authors noted that online surveys tended to yield response rates 11%-12% lower than for other types (phone, mail, and face-to-face interviews). Nevertheless, online surveys have the advantage of being more cost-effective due to the reduction in paper and postage and the reduction in time for hand coding of data or the purchase of optical character reading software or hardware necessary to create a database from paper surveys.

Some researchers have also identified unique challenges in surveying police officers, which can potentially lower the response rates even more than for other groups.[36] These concerns

[34] Three were from economic and social sciences, and the other three from health surveys.
[35] The grant period began January 2020, but due to a systemwide change to a new NIJ grant tracking system, most of the tasks were delayed until funds were released later that year. As such the survey was administered during the pandemic in 2021.
[36] For example, J. Nix and colleagues (2019) described such difficulties such as a distrust of outsiders or the need to obtain

about officer surveys are not as relevant to agency-level surveys, where a single individual with access to the relevant data and information can report on the policies and practices.

Nevertheless, for detailed agency-level surveys like this eyewitness survey, time constraints in the current climate also prevent LEA administrators from being able to dedicate the resources necessary to respond to look up information such as the number of photo arrays conducted, specific policy and training details, etc. Accordingly, we did not anticipate a high response rate. Yet newer evidence has suggested that a lower response rate *may not* limit generalizability.

In fact, with the increasingly accelerated response rate declines over the last two to three decades, and partially as a result of the pandemic in early 2020, researchers across many disciplines have begun to examine the relationship between response rates and non-response bias noting little, if any, relationship (see e.g., Ash et al., 2022; Groves & Peytcheva, 2008; Hendra & Hill, 2019; Wåhlberg & Poom, 2015). Others also report that despite declining response rates, there has been no corresponding increase in non-response bias (Küfner et al., 2022). Many scientists now concur that non-response is either not correlated with, or at least is not necessarily associated with, response-bias (Czajika & Beyler, 2016; Fosnacht et al., 2017), emphasizing that sometimes surveys with very high response rates demonstrate response bias, and those with much lower response rates often do not. According to Davern (2013), "response rates lack both validity and reliability as a proxy measure of nonresponse bias" (p. 905).

Moreover, because our survey was primarily factual in nature and did not depend on

approval from supervisors and union representatives before participating. Other factors also create barriers to response. For example, the increasing number of police scholars has led to over-surveying of police officers (by scholars and their students) and survey fatigue as plausible explanations for dwindling responses. Also, with staffing shortages nationwide, there are increasing demands on officers to work longer shifts and/or overtime, creating hesitation to participate in off-duty surveys, as many agencies cannot afford to allow first responders to complete surveys on-duty when these officers are needed in the field.

opinions and attitudes, there is even less reason to suspect a response bias, as law enforcement administrators are held to a high standard for reporting factual policy and practice-related data. Recently, evidence has demonstrated that survey reliability (for opinion and attitude surveys) can be obtained even with response rates as low as 5%, depending on the sample size. For example, Fosnacht and colleagues (2017) reported that for a sampling frame of at least 500, data estimates remain reliable even with just a 5%-10% response rate. In contrast, those with smaller samples need a 20%–25% response rate to provide fairly confident estimates.

Overall Response Rate

Of the 839 surveys delivered to agencies, 187 were received for an overall response rate of 22.29%. Our achieved response rate substantially exceeds the 5%-10% range considered reliable for large samples according to Fosnacht's (2017) criteria. It is more than four times the lower parameter (5%), and more than twice the higher parameter (10%). In other words, our response rate is more than sufficient to ensure confident estimates of the broader population from which the sample was drawn, without significant concern over non-response bias. The added consideration that the survey called primarily for administrative data rather than attitudes and opinions also minimizes possible response bias.

When PERF completed their survey over 10 years ago, they obtained an overall response rate of 44%,[37] about twice that of the current survey. Our lower response rate is likely due to the national trend of declining survey participation, compounded by the specific impact of the Covid-19 pandemic (see Ash et al., 2022).

The response rates to this survey were calculated by the number of responses from the

[37] Their sample was 1,401, but they excluded the 24 agencies that did not use eyewitness identification in calculating the response rate, increasing the rate by 1%; however, we included all respondents even if they did not conduct eyewitness identifications.

total number of agencies sampled, or in the case of state agencies, the entire population of state police, highway patrol, and investigative units from the NDLEA. For the stratified, random sample, we summarize response rates by the strata relied upon for our sampling: agency type, size (number of sworn personnel), and region. Since these strata represent smaller groups within the full respondent group, caution should be taken when interpreting the findings for any one group, for example, small agencies in a single region, etc.

It is important to note that among the 187 respondents, 139 (74%) came from agencies that were presented with the entire survey, and 48 (26%) came from agencies that completed a condensed version[38] in which just 10 key questions were presented to reduce agency burden and increase the response rate.

Actual Response Rate by Agency Type

The highest response rate was from state police, highway patrol, and criminal investigative bureaus combined (42%), followed by those from local (city, town, and county) police departments (25%), as shown in **Table 4**. Compared to the findings from over a decade ago, our response rates were lower, except when considering state agencies, for which our response rate of 42% almost mirrored that of PERF's at 43%. The fact that sheriffs' offices responded at the lowest rate (16%) was consistent with the earlier results. The lower response rate for sheriffs is likely attributable to the same reason that PERF cited in their 2013 report, rather than an unwillingness to participate: not all sheriffs' offices or departments provide investigative functions in the field because some are primarily responsible for jail management, civil process, and/or court security. In those cases, it is likely that many did not consider this survey applicable to the functions they perform.

[38] Sent after the last reminder, requesting responses to the 10 key questions deemed of most interest to the research team.

The larger response rate obtained from the smallest group (states, n = 74) suggests that despite the limited number, it is still likely to be reliable. Fosnacht (2017) recommended a minimum response rate of 20%-25% for samples under 500, and we achieved a considerably higher rate of nearly 42%. Because our sample of police and sheriffs combined was large (n = 765), the combined 20.39% response rate is well above the rates considered sufficient (5%-10%) for reducing response bias and ensuring reliable estimates.

Table 4. *Response rate by agency type (unweighted)*

Agency Type	Frequency	# Surveyed	Response Rate	2013 Rate
State Police, Highway Patrol or Investigative Bureau	31	74	41.9%	42.9%
Local Police	102	417	24.5%	51.5%
Sheriff's Office	54	348	15.5%	36.5%
Total	**187**	**839**	**22.29**	**45%**

Response Rate by Agency Size (number of sworn personnel)

Response rates increased with agency size (see **Table 5**). Our agency size groups consisted of the following: The "small" category was made up of agencies we classified as the smallest (less than 25 sworn personnel), very small (25-49 sworn personnel), and small (50-99 sworn personnel), whereas the "large" agency category consisted of large (500-999 sworn) and the largest (1,000+ sworn) agencies.

As summarized in **Table 5**, the response rate for agencies classified as "small" (< 100 sworn personnel) was just 19%, whereas the medium and large were considerably higher (30% and 32%, respectively). The same pattern held across all groups, with the smallest, very small, and small agencies yielding response rates of 16%, 18%, and 23%, respectively, and the largest group yielding a response rate of 41% compared to the large agencies' rate of just under 27%.

Table 5. Response rate by agency size (unweighted), excludes state agencies [a]

Sworn Number	# Responded	# Surveyed	Response Rate	2013 Rate
Small[b] (< 100)	91	481	18.9%	41.5%
Medium (100 – 499)	48	160	30.0%	58.4%
Large[c] (500 +)	15	47	31.9%	80.7%
Unknown[d]	2	77	2.6%	--
Total	**156**	**765**	**20.4%**	--

[a] State agencies: 31 respondents,74 total in sample (data on sworn agency size was not available)

[b] Small consists of smallest, very small, and small. [c] Large consists of large and largest.

[d] The agency size was unknown at the time of distribution.

Response Rate by Geographic Region

We relied upon the same geographic regions that PERF (2013) used, and consistent with their findings, the response rate for agencies in the Northeast was the lowest. As shown in **Table 6,** the response rate for agencies in the Northeast was 15%, whereas those from the West, Midwest, and South responded at a higher rate (24%, 22%, and 20% respectively). However, when considering just the state agencies (see **Table 7**), a different picture emerged altogether, with the Northeast having the highest response rate (60%), followed by the Midwest (53%) and the South (50%), whereas those in the West had a markedly lower rate of just 14%.

Table 6. Response rate by geographic region (unweighted), excludes state agencies [a]

Region	Local Police & Sheriff	# Surveyed	Response Rate	2013 Rate
West	46	195	23.6%	50.1%
Midwest	47	212	22.2%	40.3%
South	36	178	20.2%	49.7%
Northeast	27	180	15.00%	38.5%
Total	**156**	**765**	**20.4%**	**44.7%**

[a] State agencies: 31 respondents of 74 surveyed

Table 7. Response rate for state agencies by geographic region (unweighted)

Region	States	# Surveyed	Response Rate	2013 Rate
West	3	21	14.3%	n/a
Midwest	9	17	52.9%	n/a
South	13	26	50.0%	n/a
Northeast	6	10	60.0%	n/a
Total	**31**	**74**	**41.9%**	**42.9%**

See also **Appendix D**: Map of Agency Survey Responses by Region.

Survey Responses

From this section forward, the reported response frequencies (*n*) in all tables represent the *weighted* frequency based on our stratified sample. However, for state agencies, the frequencies reflect the *actual* obtained values, as we surveyed all state agencies rather than sampling. Also, it is important to note that over 25% of respondents were only presented with 10 key questions, so for many questions, the number of respondents is based on 139 participants rather than 187. Finally, most questions related to the year-long period ending December 31, 2019. This is important because the pandemic, identified early in 2020, led to significant anomalies, such as irregularities and inconsistencies in law enforcement activity and data (Lum et al., 2023; Nielson et al., 2022).

Characteristics of Responding Agencies

Characteristics reported here differ from the response rate data for two reasons: (1) they are based on the responses provided by the agencies, not the NDLEA data, and (2) the actual response rate section does not use weighted values. Among the 187 responding agencies, 72% reported that they were from local police departments, 12% from sheriffs' departments or offices, and 16% from state police, highway patrol, or investigative bureaus. In terms of the regions reported, the greatest proportion was from the Midwest (36%), followed by the South

(33%), the Northeast (20%), and the lowest proportion was from the West (11%).

For agencies reporting the population they served (n = 107[39]), just 7% served populations of 1,000-4,999, whereas 34% served populations between 5,000 and 9,999. Another 29% worked in communities with populations between 10,000 and 99,999, and 15% did so in jurisdictions with 100,000-999,999. Finally, 15% of respondents served populations of a million or more, with most having at least 2 million.

For those agencies reporting their number of sworn personnel (n = 186), over half had less than 25 sworn (57%), with 20% having 25-99 sworn, 15% with 100-499, and the remaining 9% with 600 or more. When considering the number of civilian personnel (n = 114), 15% had none, whereas almost half had less than 25 (47%). Another 17% had 25-99 civilian employees, and the remaining 21% had 100 or more.

The majority of those individuals who completed the survey indicated they were sworn (99%). With respect to the respondents' rank, 82% were members of the command staff or top leadership, and the remainder were mostly sergeants (12%), a few detectives (4%), and just one line officer.

Responses to Key Questions

Eyewitness Procedure Types Used

Across agency types, 86% of 157 agencies reported conducting photo arrays, and 64% reported conducting live showups. Compared to the 2013 PERF findings, respondents were somewhat less likely to conduct photo arrays (86% vs. 94%), although this was not the case for showups (64% vs. 62%). If agencies reported that they did not conduct showups, we asked why

[39] Excludes six agencies whose responses were "not applicable" (presumably from state agencies).

they did not. Over two dozen agencies included a response. Among them, four agencies responded that their local district attorney, prosecutor, or county attorney prefers they be avoided, although one respondent noted that they are still allowed. A handful of other respondents indicated that they were not necessary (no opportunity) or that they do not conduct criminal investigations within their agency.

Types of Eyewitness Procedures by Agency Type. As shown in **Table 8**, a lower proportion of local police reported conducting photo arrays (84%) than sheriffs or state agencies (91% and 90%, respectively). Yet, a considerably larger proportion of local police reported conducting live showups (74%) as compared to sheriffs' or state agencies (29% and 48%), as detailed in **Table 9**.

Table 8. Number of agencies that conduct **photo arrays** by agency type

	Local Police	Sheriff's Office or Department	State Agency	**Total**
Yes	90 (84.1%)	19 (90.5%)	26 (89.7%)	135 (86%)
for investigative purposes only	*22 (20.6%)*	*5 (23.8%)*	*5 (17.2%)*	*32 (20.4%)*
for investigative purposes & to submit to the court as evidence	*68 (63.6%)*	*14 (66.7%)*	*21 (72.4%)*	*103 (65.6%)*
No	17 (15.9%)	2 (9.5%)	3 (10.3%)	22 (14%)
Total	107 (100%)	21 (100%)	29 (100%)	157 (100%)

Table 9. Number of agencies that conduct **live showups** by agency type

	Local Police	Sheriff's Office or Department	State Agency	**Total**
Yes	86 (73.5%)	6 (28.6%)	14 (48.3%)	106 (63.5%)
for investigative purposes only	*33 (28.2%)*	*2 (9.5%)*	*4 (13.8%)*	*39 (23.4%)*
for investigative purposes & to submit to the court as evidence	*53 (45.3%)*	*4 (19.0%)*	*10 (34.5%)*	*67 (40.1%)*
No	31 (26.5%)	15 (71.4%)	15 (51.7%)	61 (36.5%)
Total	117 (100%)	21 (100%)	29 (100%)	167 (100%)

Types of Eyewitness Procedures by Agency Region. As demonstrated in **Table 10**, agencies in the Midwest, Northeast, and West were significantly more likely to conduct photo arrays (all above 93%) than were agencies in the South (about 62%). A similar pattern was true for showups, with between 63% and 80% of those agencies reporting they conduct showups as compared to just 36% of agencies in the South (see **Table 11**).

Table 10. Number of agencies that conduct **photo arrays** by agency region

Response	Midwest	Northeast	West	South	**Total**
Yes	55 (93.2%)	35 (97.2%)	18 (100%)	28 (62.2%)	136 (86.1%)
No	4 (6.8%)	1 (2.8%)	0	17 (37.8%)	22 (13.9%)
Total	59 (100%)	36 (100%)	18 (100%)	45 (100%)	158 (100%)

Table 11. Number of agencies that conduct **showups** by agency region

Response	Midwest	Northeast	West	South	**Total**
Yes	53 (80.3%)	24 (63.2%)	14 (70%)	16 (36.4%)	107 (86.1%)
No	13 (19.7%)	14 (36.8%)	6 (30%)	28 (63.6%)	61 (13.9%)
Total	66 (100%)	38 (100%)	20 (100%)	44 (100%)	168 (100%)

Types of Eyewitness Procedures by Agency Size. The proportion of agencies conducting photo arrays grew with agency size, with all large agencies conducting them, whereas small and medium agencies were somewhat less likely to do so (see **Table 12**).

Table 12. Number of agencies that conduct **photo arrays** by agency size

Response	Small (< 100)	Medium (100 – 499)	Large (500 +)	**Total**
Yes	94 (83.2%)	23 (88.5%)	18 (100%)	135 (86%)
No	19 (16.8%)	3 (11.5%)	0	22 (14%)
Total	113 (100%)	26 (100%)	18 (100%)	157 (100%)

On the other hand, when it came to showups, a smaller proportion of large agencies reported conducting them (56%) as compared to small (62%) or mid-sized (73%) agencies, see **Table 13.**

Table 13. Number of agencies that conduct **showups** by agency size

Response	Small (<100)	Medium (100 – 499)	Large (500 +)	**Total**
Yes	76 (62.3%)	19 (73.1%)	10 (55.6%)	105 (63.3%)
No	46 (37.7%)	7 (26.9%)	8 (44.4%)	61 (36.7%)
Total	122 (100%)	26 (100%)	18 (100%)	166 (100%)

Confidence Statements Obtained in Eyewitness Procedures

During the past decade, more evidence has accumulated supporting the strong relationship between initial confidence (during the first procedure) and accuracy of the identification (Arndorfer & Charman, 2022; Wixted & Wells, 2017).[40] Specifically, only confidence statements collected right after the first identification procedure and/or decision matter, and all later statements come at a cost of memory loss and reduced accuracy. Moreover, some have asserted that this relationship is only strong when the photo array is conducted in an unbiased manner (Wixted & G. Wells, 2017), or what is referred to as "pristine" testing conditions, although others have argued that such conditions may not be necessary for the relationship to remain strong (see Mickes et. al., 2017). We asked agencies to report on whether they request confidence statements from witnesses in both showups and photo arrays, and if they were aware of this newer research that confidence expressed during the initial viewing is highly predictive of accuracy.

Agencies Collecting Confidence Statements. For showups, 72.3% of agencies reported they collected confidence/certainty statements just after completing the procedure,[41] and 84% said they did so for photo arrays. Displayed in **Table 14** are the proportions of agencies collecting

[40] But also see Moore et al. (2024)

[41] We captured confidence/certainty slightly differently than PERF, such that this response is based on the agencies for which the response was deduced from those that responded, "we do not capture confidence at all." We specifically asked agencies if they collect confidence statements for positive ID and nonidentifications combined, versus not capturing it at all.

confidence statements for positive IDs and non-identifications for both showups and photo arrays immediately after/during the identification decision. Compared to PERF'S earlier findings,[42] the proportion collecting confidence/certainty in showups declined from 85% to 72.3%, but the proportion for photo arrays remained largely unchanged (85% vs. 84%).

Table 14. Proportion of agencies that collect confidence statements

Response	Showups	Photo Arrays
For both positive ID and non-identification, or other[43]	73 (72.3%)	113 (84.3%)
No confidence statement	28 (27.5%)	21 (15.7%)
Total	101 (100%)	134 (100%)

Differences in Capturing Confidence by Agency Type and Size.[44] For showups, sheriffs and state agencies were *more* likely to collect confidence statements than were local police (86% and 92%[45], respectively as compared to just 68%). The same was true for photo arrays, with sheriffs and state agencies (95% and 92%, respectively) *more* likely to ask for confidence statements than local police (80%).

When considering agency size for showups, small agencies were the *least* likely to capture confidence (65%) as compared to medium (100%) and large (88%)[46] ones. Similarly, for photo arrays, small agencies were *least* likely to capture confidence (80%), followed by medium and large agencies (91% and 94%, respectively).

Awareness of Research Demonstrating Strong Confidence-Accuracy Relationship. Over half (54%) of the responding agencies (*n = 132*) reported they were *not* aware of the newer

[42] Whereas PERF asked about certainty for "no identification" (p. 54), we asked for certainty for "not the person."
[43] Includes those who don't ask for it if it was provided during procedure.
[44] Excludes those reporting "other."
[45] These are based on less than 15 responses per group for sheriff and state agencies.
[46] Based on < 10 responses from large agencies.

research on the strong confidence-accuracy relationship, which might explain the drop in agencies collecting confidence over the past decade.

State agencies were *least* likely to be aware of this newer evidence (27%) as compared to local police and sheriffs, who were almost twice as likely to be aware. Surprisingly, a greater proportion of respondents from small agencies were knowledgeable of this new evidence on confidence and accuracy (54%) as compared to large (29%) and medium-sized agencies (23%). This was the case despite these small agencies being the least likely to collect confidence statements, as reported above.

How Confidence is Expressed. For showups, 77% of the responding agencies said they request the confidence statement in one's own words, while 15% ask for it on a numeric scale (1-10, 1%-100%, etc.), 2% on a verbal scale (e.g. not at all certain to absolutely certain, not sure to definitely the person, etc.). The remaining 6% reported requesting the confidence statement in a different way.[47] For photo arrays, 74% request that witnesses express confidence in their own words while 13% of agencies each ask for confidence using either a numerical or verbal scale.[48]

Key Questions Specifically Related to Photo Arrays

One area of great research attention over the past decade is the reliability and diagnosticity of identifications made using traditional, simultaneous lineups (typically 6 or more photos shown concurrently) versus sequential (where photos are presented one at a time). While some researchers previously indicated that a sequential presentation format was superior to the traditional format (see Lindsay & G. Wells, 1985; Steblay, et al., 2001, 2011), other researchers

[47] Seventeen agencies described the specific numeric scale used, with 13 using a percentage scale (76%), 3 using a 1-10 scale and one other reporting that either was acceptable.
[48] Twenty-nine agencies wrote in the specific scale used with most (70%) using a percentage scale, six using a 1-10 scale, and two others said it could be either.

concluded that simultaneous procedures were at least as reliable and/or diagnostic as sequential procedures (Amendola & Wixted, 2015a, 2015b; Gronlund et al., 2009; Gronlund et al., 2014; Mickes et al., 2012), leading the National Research Council (NRC) to suggest the issue was not fully resolved (2014).

According to the 10-year-old NRC Report on eyewitness identification, "there is no consensus among law enforcement professionals as to whether the sequential presentation procedure is superior to the simultaneous procedure" (p. 54). However, since then, more research has accumulated demonstrating that, if anything, the simultaneous method may in fact be superior to the sequential one (Wixted et al., 2016; Mickes & Wixted, 2025). To gain further insight into this evolving research, we asked respondents which method they used or if they used both methods.[49]

Photo Array Presentation Methods by Agency Type and Size

Overall, 61% of respondents reported using the sequential method exclusively, with just 24% relying solely on a simultaneous presentation, 2% said they used a different method, and 13% reported that they used both methods.[50] Large agencies were also most likely to use both methods (35%) as compared to medium (13%) and small (9%) agencies.

When considering whether the agencies or the states in which they reside have requirements or recommendations about presentation methods, agencies are much more likely than states to require sequential procedures (48% vs. 12%), although 35% of states recommend this practice, despite newer contradictory evidence. Moreover, 23% of agencies recommend the

[49] Agencies that allow both methods provide discretion to individual investigators or that had changed their approach in 2019, with some being done one way and others a different way during that same year.
[50] These rates were calculated after excluding three agencies who reported their method as "other".

sequential method as well. These findings are at least somewhat surprising in light of newer evidence, though not so much when considering the science-to-practice time lag. It has been over 10 years since the NRC Report noted neither method was superior, while adding that some emerging evidence suggests the superiority of the traditional, simultaneous method. Since then, more evidence has accumulated suggesting the simultaneous (6-pack type) approach is slightly better at preventing misidentifications and accurately capturing actual guilty identifications (see e.g., Amendola & Wixted, 2015b; Mickes & Wixted, 2025).

Our findings are likely due to prior evidence, which suggested a sequential superiority effect and led to recommendations by several organizations and individuals to use the sequential method. However, newer evidence and recommendations by authoritative sources no longer support a sequential superiority effect, in large part due to a less diagnostic method and associated conclusions, no longer borne out by studies using ROC analysis (refer to footnote 30 for more information). Our findings suggest that most agencies are likely not aware of the research findings over the past decade. We address this specific issue in the next section.

One other related question we asked those agencies who reported using the sequential presentation method was if witnesses were allowed to review the photos a second time ("a second lap"). Almost 82% of those agencies using sequential administration noted that they allow for a second review of the suspects.[51] About 93% of state agencies and 81% of local police allowed this, with sheriffs' agencies being somewhat less likely (73%).[52] Fully 100% of large agencies allowed a second lap, and 94% of medium agencies did so compared to just 76% of small agencies.

[51] These rates exclude responses from 5% in which the respondent selected "other."
[52] This rate is based on just 11 respondents.

Awareness of Research Demonstrating NO Sequential Superiority Effect

We queried respondents as to whether they were aware of more recent research indicating that the sequential presentation method (photos shown one at a time) may not actually be better than a simultaneous presentation (shown as a group). Just over one-third of the 132 responding agencies (34%) reported that they were aware. Slightly more local police were aware of the lack of a sequential superiority effect (37%) compared to sheriffs and state agencies (both about 28%). Also, when considering agency size, the smallest agencies were more likely to be knowledgeable about this research (39%) than large (31%) and mid-sized agencies (18%). In the narrative comments provided by agencies, a few said that they believed it to be just the opposite, likely because the state of the science well over a decade ago indicated a likely sequential advantage. We believe this to be indicative of the science-to-practice lag, in part due to limited dissemination to the decision-makers in LEAs and the slow pace of adoption.

Responses to Non-Key Questions

Showups and Photo Arrays

Policies. One of the key findings in the PERF report was that many agencies had not adopted policies for photo arrays or showups. Specifically, 64% of agencies had no written policies for photo arrays, and 77% had no written policies for showups. In the current study, we found a substantially increased proportion of agencies with formal written procedures for both types (see **Table 15**). Only 8% of the respondent pool indicated that they do not have a formal policy for photo arrays and for showups. With respect to photo arrays, 92% reported that their agencies have policies for both constructing and administering the procedure.

Policy Mandates at the Local and State Levels. In this section, we outline required or recommended practices at both the state and agency levels. In **Table 16**, we present state and

Table 15. Proportion of agencies having **no** written policy

Type of procedure	NPI 2021 survey	PERF 2011 survey
Photo array	92% *(n = 73)*	36% *(n = 584)*
Show-up	92% *(n = 73)*	23% *(n = 580)*

agency-level requirements and recommendations related to showup procedures. We also note those for which there is neither a requirement nor a recommendation, or in other words, the state or agency was silent on this issue at the time of the survey. This is followed by the same procedures for photo arrays, also by state and local policies (see **Table 17**).

Table 16. State and agency requirements and recommendations on **showups**

Item	*State n*	*Agency n*	State			Agency		
			REQ	REC	S/UNK	REQ	REC	S/UNK
Video record procedure	74	74	12%	47%	41%	53%	28%	19%
Audio record procedure	74	75	12%	47%	40%	52%	31%	17%
Written report by administrator of procedure	75	75	28%	37%	35%	73%	7%	20%
Written report by someone other than administrator	74	74	7%	32%	61%	18%	18%	64%
Computerized administration	73	71	1%	29%	70%	11%	16%	74%

Table note. REQ = requires; REC = recommends; S/UNK = silent or unknown

Table 17. State and agency requirements and recommendations on **photo arrays**

Item	State & Agency n	State			Agency		
		REQ	REC	S/UNK	REQ	REC	S/UNK
Video record procedure	84	16%	46%	38%	48%	23%	30%
Audio record procedure	85	17%	45%	38%	47%	22%	31%
Written report by administrator of procedure	85	38%	23%	39%	91%	2%	7%
Written report by someone other than administrator	85	21%	30%	49%	54%	8%	38%
Computerized administration	84	1%	28%	71%	17%	19%	64%
Sequential presentation	84	12%	35%	52%	45%	29%	26%

Instructions to Witnesses in Identification Procedures. This section provides information about whether standardized instructions are provided to witnesses, the format of the instructions, whether or how they are documented, and the degree to which agencies have included some recently recommended instructions.

As shown in **Table 18**, the most common approach was to use standardized, *oral* instructions, with 73% using this method for showups and 85% for photo arrays. More than half of respondents reported providing standardized, *written* instructions in photo arrays (56%), though in showups, the proportion was significantly lower. These response categories are not mutually exclusive such that some agencies may provide instructions in more than one manner.

Table 18. Means for conveying instructions to witnesses

Instruction delivery method	**Showups** *n*	*n respondents*	**Photo Arrays** *n*	*n respondents*
Oral/verbal (standardized)	46 (73%)	*63*	73 (85%)	*86*
Written (standardized)	13 (21%)	*62*	47 (56%)	*84*
Oral/verbal (informal)	33 (54%)	*61*	41 (49%)	*84*
Oral recorded instructions	35 (56%)	*63*	35 (41%)	*85*

For showups, 81% reported requiring that instructions provided to witnesses be documented[53] compared to 95% of those conducting photo arrays. Also, 81% reported requiring documentation of how the instructions are provided to the witness for showups[54] and 93% for photo arrays. Almost all agencies reported requiring the administrator to document whether the suspect was selected (99% for showups[55] and 100% for photo arrays).

For photo arrays, slightly fewer agencies (94%) required documentation of a filler identification, as opposed to a positive ID of the suspect. Finally, 99% of respondents required

[53] 4% said it was discretionary.
[54] 5% said it was discretionary.
[55] 1% said it was discretionary.

documentation of the confidence statements for positive IDs, but just 86% when an ID was not made. Another notable finding was that 94% of agencies responded that they require documentation of the reason a person was identified as a suspect, and/or was included in the photo array.

We also asked agencies whether they included specific recommended instructions provided in various guidelines. Detailed in **Table 19** are the number and proportion of agencies that include each of seven specific instructions broken down for both showups (where applicable) and photo arrays.

Table 19. Detailed instructions provided in showups and photo arrays

Instruction Provided	**Showups** *n (62 – 63)*	**Photo Arrays** *n (84-86)*
You do not need to make an identification	59 (94%)	75 (89%)
The individual you will see may/may not be the person who committed the crime (showup)	60 (97%)	n/a
OR		
The perpetrator may/may not be in the photo array	n/a	74* (86%)
The suspect's appearance may have changed since the crime	n/a	67 (79%)
The investigation will continue with/without an ID	57 (90%)	65 (77%)
It is just as important to avoid identifying an innocent person as it is to identify a guilty person	n/a	61 (71%)
The agency cannot provide any feedback about their decision	n/a	46 (54%)
Do not search for a photo of the perpetrator online	n/a	30 (35%)

As shown in **Table 19**, almost all agencies conducting showups indicate that the witness does not have to make an identification (94%), that the individual the witness will see may or may not be the person who committed the crime (97%), and that the investigation will continue whether or not they make an id (90%). In comparison, the proportion of agencies doing so for photo arrays was somewhat lower. For photo arrays, we also asked whether agencies advise witnesses that the perpetrator's appearance may have changed since the crime and that clearing

an innocent person is just as important as identifying a guilty one; 79% of agencies do so for the former and 71% for the latter.

In **Table 20**, we compare the provision of these instructions in our study with those in the PERF study conducted over a decade ago. For both photo arrays and showups, there were marked increases in the proportions of agencies instructing witnesses that they do not need to make an identification, and that the investigation will continue even if an identification is not made. This was also the case for instructing witnesses that the suspect's appearance may have changed in the case of photo arrays. The instruction that the person viewed may or may not be the perpetrator has long been a strong caution in showups, with over 90% previously including it. The fact that about 15% of agencies do not provide this same instruction for photo arrays stands out. Notably, it has not gone up significantly since the PERF survey.

Table 20. Comparison of witness instructions in current NPI and prior PERF survey

Instruction Provided	**NPI** ***Showups***	**PERF** ***Showups***	**NPI** ***Photo Arrays***	**PERF** ***Photo Arrays***
You do not need to make an identification	94%	62%	89%	56%
The individual you will see may/may not be the person who committed the crime *OR*	97%	94%	--	--
The perpetrator may/may not be in the photo array	--	--	86%	84%
The suspect's appearance may have changed since the crime	--	53%	79%	60%
The investigation will continue with/without an ID	90%	62%	77%	60%

Acknowledgment of Instructions. Agencies generally require that witnesses acknowledge the instructions provided to them, with 85% requiring this for showups and 95% for photo arrays. For showups, almost an equal proportion required verbal (49%) and written (47%) acknowledgment, with just 4% requiring it in some other manner. For photo arrays,

however, most required a written acknowledgment (63%), with less than a third requiring a verbal one (30%), and just 2% in some other way. Additionally, 96% of those requiring a written acknowledgment in photo arrays also required a signature.

Instructions to Officers. In this section, we provide data on whether agencies have policies on how to administer and document eyewitness procedures. About 95% of agencies provide instructions on *administering* photo arrays, while 92% do so for showups. For photo arrays and showups, 93% said officers were provided with instructions on how to *document* the procedures.

Latency, or the time to make an identification, is an important variable in eyewitness identification because the research evidence suggests that identifications made more quickly are more likely to be accurate (Carlson, et al., 2025; Dobolyi & Dodson, 2018, Semmler et al., 2018), For photo arrays, 66% capture latency as a result of using an audio/video recording, whereas 34% do not capture latency. For showups, 67% of 66 agencies that capture latency document by recording the procedure, while 27% do not.

Personnel Authorized to Conduct Showups and Photo Arrays. We asked whether the agencies had specified (through policy or procedure) which personnel are allowed to conduct an eyewitness identification procedure. For showups, 47% of 74 responding agencies noted that their agencies did specify who is allowed to conduct a showup, whereas for photo arrays 69% of 86 responding agencies did so.[56]

Showups. For showups, 100% of responding agencies allowed patrol officers/deputies, detectives/investigators, and sworn supervisors/commanders to conduct them, while about 93%

[56] A considerably higher proportion of local police (86%) specified who is authorized to conduct a photo array as compared to sheriffs' or state agencies with 23% and 57% respectively. By agency size, large agencies were least likely to specify this.

permitted other sworn officers to do so.

Photo Arrays. For photo arrays, 93% of agencies authorized patrol officers/deputies to conduct the procedure, and 86% authorized detectives/investigators, regardless of their assignment to the case. Additionally, 97% of agencies authorized supervisors or commanders to conduct photo arrays, while only 25% authorized other sworn officers, with 21% being unsure. Only one agency allowed a civilian to conduct a photo array.

Training for Officers. Officers in 85% of responding agencies are provided with training in conducting showups, and 83% are provided with training in photo arrays. For photo arrays, local police agencies are more likely to provide training to their personnel (94%) than are sheriffs (50%) or state agencies (74%), and medium and large agencies are slightly more likely to provide training (86% for both) than are smaller ones (80%). For showups, there were too few respondents in some categories to provide generalizations about training provided by agency type and size.

Type of Training Provided. For showups, over two-thirds of agencies provide a combination of formal and/or on-the-job training to their officers, with just over 70% that provide it to detectives. For photo arrays, almost 83% of agencies provide a combination of formal and/or on-the-job training to their personnel. For patrol officers/deputies as well as sworn supervisors, 87% are provided with training, and for detectives, 89%. For 88% of agencies, they either do not provide or do not know if their agency provides training to civilian employees.

Who Provides Training? Most agencies we surveyed responded that their own agency provides training to their officers on eyewitness identification practices (97% for both types), whereas in 2013, just 50% of agencies said they provided their own training. **Table 21** shows the proportions of agencies that use each type of training provider for formal/informal training.

Table 21. Proportions of agencies using various training providers

Training Provider	Administering Showups *(n = 61-63)*	Conducting Photo Arrays *(n = 68-71)*
Agency Itself	97%	97%
Prosecutor/District Attorney	43%	50% [a]
State Law Enforcement Agency (LEA)	26%	25% [b]
Other LEA	24%	--
OR County LEA	--	12% [b]
Federal LEA	6%	6% [b]
Psychologist/Eyewitness Scientist	3%	4% [b]

[a] 21% reported they did not know

[b] 45 – 50% reported they did not know

Feedback. One important piece of guidance designed to reduce bias in eyewitness procedures is to prohibit officers from providing any feedback to witnesses after their identification procedure is completed. For photo arrays, 63% of agencies said they prohibit officers from providing feedback. For showups, 67% of respondents said they prohibit this practice, whereas 12% did not know.

Changes made to eyewitness procedures over the last decade. For both showups and photo arrays, we asked responding agencies whether they had changed their procedures over the last decade. For photo arrays, 77% said they had, 14% said they had not, and the remaining 8% were not sure. For showups, 69% said they had changed policies in the past 10 years, whereas 19% said they had not done so, and 12% were unsure.

Showups. Among agencies that made changes in the past decade, 57% reported changing how they administered showups, and 51% reported changing how they documented them. For both types of changes, the agencies were about evenly split between making those changes in the past three years and doing so more than three years ago.

Photo arrays. Of the 69 agencies reporting they had made changes, 72% reported changes in how photo arrays were administered, with almost three-fourths of those having made changes more than three years ago, and about a fourth reporting they made changes within the past three years. Fully 71% of respondents reported having changed their witness instructions.

Moreover, 46% noted that they changed procedures related to filler selection, number, and their placement in the lineup. Regarding the selection of fillers, 38% of 71 respondents noted that they began using electronic devices to select fillers during the past decade, whereas 20% indicated that they had been doing that before 2010. Similarly, for the administration of photo arrays, 24% noted that they had begun using electronic devices to conduct photo arrays in the past decade, whereas 10% claimed that they had already done so before 2010.

Among 68 respondents, 51% indicated they had changed the procedure for documenting photo arrays in the past decade, 63% made changes over three years ago, and 37% made those changes in the past three years.

Source of Information Relied Upon to Establish or Modify Policies/Practices. Agencies were asked to indicate which sources they relied upon during the past decade to establish or modify policies and practices. This allowed us to assess which type of sources LEAs see as most authoritative on matters of eyewitness identification practice or policy. The sources are broken down for showups and photo arrays (see **Table 22**).

Table 22 includes information not previously collected from LEAs and points to several important aspects of policy formation and modification regarding eyewitness identification procedures as well as the research-practice gap discussed earlier in the report. First and foremost, prosecutors or local district attorneys appear to be most influential in shaping agency policies, with 71% of agencies relying upon them for showup practices and 68% for photo arrays.

Following closely are the professional associations to which agencies belong, such as the IACP for police departments, National Sheriff's Association (NSA) for sheriffs, the International Association of Directors of Law Enforcement Standards and Training (IADLEST), which is made up of state Peace Officer Standards and Training Commissions, and others such as state or regional associations and the like with 70% for showups, though clearly less so for photo arrays. Among the most common sources are also state regulations and/or recommendations like those provided in state attorneys' general guidelines. Many also wrote in that they relied on the Commission on Accreditation for Law Enforcement Agencies (CALEA) as an authoritative source.

Table 22. Sources consulted for establishing/modifying eyewitness policies and practices

Source	**Showups** *n = 66*[a]			**Photo arrays** *n = 132*[b]		
	Yes	**No**	**Unknown**	**Yes**	**No**	**Unknown**
Adopted a policy from another LEA	n/a	n/a	n/a	33%	44%	24%
Obtained advice from another LEA	55%	24%	21%	24%	57%	19%
Professional association (*e.g., IACP, NSA, IADLEST, etc.*)	70%	14%	17%	53%	14%	33%
State required/requested or State Attorney General Guidelines	64%	21%	15%	48%	33%	20%
District Attorney/Prosecutor	71%	20%	9%	68%	15%	17%
Advocacy Group (e.g., Innocence Project, ACLU, etc.)	8%	58%	35%	14%	48%	38%
Advised by Psychologist or Social Scientist	24%	53%	23%	10%	60%	30%
U.S. Dept. of Justice (Yates memo)	11%	23%	67%	12%	39%	49%
2014 National Research Council Report	5%	29%	67%	4%	42%	54%
1999 National Institute of Justice Guidelines (U.S. Department of Justice)	11%	36%	53%	11%	43%	46%
Reviewed some scientific publications	36%	29%	35%	22%	47%	31%
Relied upon case law	38%	30%	33%	26%	25%	48%

[a] *Case law n = 64*

[b] *Psychologist n = 131;Case law n = 114*

A few other sources of guidance stood out as somewhat remarkable. First, it is notable that one-fourth to as many as 38% of LEAs rely on case law in this area, especially in light of the assertions pointed out earlier in this report that court decisions often do not reflect the most up-to-date scientific evidence. Next, it appears that LEAs are getting more sophisticated in examining scientific findings, with 22 – 36% reviewing scientific publications. We are unable to assess from the survey how agencies weigh the quality and rigor of such research in making policy/practice decisions. Also, while psychologists and other social science researchers have produced a large evidence base regarding aspects of eyewitness memory and eyewitness identification practices, a small proportion of agencies (10 % for photo arrays, and 24% for showups) rely on guidance from these researchers.

Finally, considering the research-practice lag, it is perhaps not surprising that few agencies have relied upon the 2014 report of the National Research Council. It is unclear whether this is due to a lack of knowledge that the report was issued, its length, complexity, or simply the slow pace of dissemination and diffusion. At the same time, it is possible that the NRC Report may have informed newer guidance by CALEA or IACP. Nevertheless, some have overstated the influence the NRC Report has had on field-based practice. For instance, while Albright & Garrett (2022) have asserted that it has shifted practices remarkably, there does not appear to be clear evidence of its impact per se, and our survey results do not allow us to draw that conclusion. Other recommendations may also be the reason for some shift in practices. Some of the changes since the time of the PERF survey appear to be guided more by district attorneys, prosecutors, or states, as well as guidance and policy recommendations and standards by professional associations, such as IACP and CALEA, both of which have updated their standards

since the NRC Report. The extent to which the NRC Report findings has influenced those organizations is unknown.

Questions Related Specifically to Photo Arrays

Number of Photos, Suspects, and Fillers Captured in Photo Arrays. Most agencies include six photos in each array (92%), with 7% displaying eight and 1% showing nine. No agencies used seven photos or more than nine.[57] Also, 88% of respondents said they had a policy for how many suspects are allowed to be included in a photo array procedure,[58] and most said they only allowed one suspect per lineup procedure (88%), with 11% reporting that there was not a clear policy about this. Regarding the number of fillers, 78% of agencies allow for five to six fillers in a photo array, whereas just 8% each allowed for just three to four or more than seven. About 6% said they had no clear policy or responded "other."[59]

Position of Suspect in Photo Arrays. When asked if there is a position in a photo array in which a suspect is never placed, 74% said no. However, 17% noted that they never place the suspect in position one, and 8% said never in position six. Among 85 responding agencies, 53% have the person constructing the photo array determine the order of the suspect/fillers, while 32% allow the person administering the photo array to determine that order. Just 1% allow the suspects and/or their attorneys to select the order.

Use of Computer/Software for Creating Photo Arrays. Over two-thirds of the 86 agencies that responded noted that their agency does use software to create their photo arrays (67%). Of those, about 16% developed their own program, 28% purchased or leased the software

[57] Local police and sheriffs were slightly more likely than state agencies to rely on six photos (96% and 92% versus 83%), as were small and large agencies (94 and 93%) as compared to mid-sized ones (86%).
[58] A greater proportion of local agencies had such a policy (98%) as compared to state agencies (83%) and sheriffs' agencies (73%). Also, all medium-sized agencies had a policy on number of suspects as compared to small (90%) and large ones (80%).
[59] Local police were more likely to use 5-6 fillers (87%) as compared to sheriffs' or state agencies (77% and 70% respectively). Small and large agencies were more likely to use 5-6 fillers (86% and 80%) as compared to mid-sized agencies (52%).

from a vendor, and almost 57% acquired it another way. Among 84 respondents, 52% said they use software to randomly select the order of photos in photo arrays as well.

Guidance Provided to Administrators in Photo Arrays. Agencies provide guidance, either through training or policy, on several recommended practices for administrators of photo arrays to prevent bias or suggestibility.

Blinding. One form of guidance is the use of a double-blind or blinded procedure to eliminate the possibility of non-verbal cues by officers, even when unintentional, that may influence the witness in the procedure. The primary way to ensure this is to have a different person administer the photo array than the one who constructed it, such that the administrator is unaware of which person is the suspect ("double-blind"). However, it can also be achieved even if the administrator is aware of who the suspect is, as long as they stay out of view of the witness so as to eliminate the possibility of non-verbal cues ("blinded"). From some of the comments provided by responding agency personnel, it became apparent that some conflate presentation method (sequential vs. simultaneous) with blinding procedures. This may be because of the way it has been described by various scientists, experts, or trainers, or by virtue of the way the individual policy is written. Much of this came from some earlier recommendations that agencies use double-blind sequential procedures. But double-blinded simultaneous procedures are also possible, as are unblinded sequential procedures.

Part of the confusion may arise when considering the sequential procedure, where blinding can only be achieved if the administrator is unaware of the order of photos, since a decision is required after each individual photo is shown. This can be achieved in one of two ways, depending on whether the photo array is administered electronically or with paper photos. For electronic presentation, the order can be randomly assigned by the computer software, and as

long as the administrator did not review the order, and cannot see the screen, this will create the blinded condition. For a paper or non-electronic presentation using a sequential method, the 'folder shuffle' method is recommended. This method involves placing suspect and filler photos into separate folders and then providing them to the administrator to shuffle before presenting the folders, one at a time, to the witness. In this case, the administrator has no idea which photo is in which folder (double-blinded).

Surveyed agencies were asked whether their photo array administrators are asked not to sit or stand next to the witness or in a position where they can see what the witness sees. 79% said they followed this approach. Also, 62% of agencies that reported using a paper method for sequential photo arrays indicated that they advised photo array administrators to use the folder shuffle method. Moreover, 55% of 107 agency respondents reported that the case detective knew which photo was the suspect at the time of the procedure. Just 27% said the lineup administrator knew the suspect's position in the array. Far fewer (5%) said the district attorney or prosecutor knew the position, and importantly, 13% said none of these individuals were aware of which photo was the suspect at the time of the administration.[60] The latter finding indicates that for those 13%, they are likely using a double-blind administration procedure.

Pre-administration questions. Based on some recommended practices, agencies were queried about information obtained from witnesses prior to administering photo arrays. **Table 23** lists the items required to be gathered or documented prior to the procedure.

Guidance on Administration, Creation, and Documentation of Photo Arrays. We were interested in examining whether specific guidance has been provided either through training or policy. These specific issues and guidance are provided in **Table 24,** presented in order of most

[60] Excludes 11 who said "other."

common to least prevalent guidance.

Table 23. Instructions to witnesses prior to photo array procedures

Questions Answered and/or Documented	Yes
How far away from the perpetrator they were from perpetrator or if something obstructed their view (n = 84)	64%
The amount of time that the witness saw the perpetrator during the crime (n = 85)	64%
If the perpetrator had a weapon (n = 83)	61%
If someone showed them a photo of the suspect prior to this procedure (n = 84)	43%
If the witness searched social media/internet and/or believes they found the culprit (n = 84)	41%
If the suspect's photo has been made public (n = 84)	38%

Table 24. Information provided in training of agency personnel in photo arrays

Content	Frequency
Avoid saying anything that may influence the witness' selection *(n = 86)*	83 (97%)
Preserve photos used in every array that is presented to any witness(es) *(n = 69)*	67 (97%)
Avoid reporting any information about the person the witness selected prior to obtaining a confidence statement *(n = 69)*	66 (96%)
Fillers should generally fit the witness' description of the perpetrator *(n = 68)*	65 (96%)
Advise witness not to discuss the procedure or results with other witnesses *(n = 86)*	78 (91%)
At least five fillers should be used in each photo array *(n = 69)*	67 (91%)
Preserve presentation order of photos *(n = 69)*	63 (91%)
Place suspects in different positions *(n = 70)*	60 (86%)
Avoid providing any feedback to the witness after their participation *(n = 85)*	73 (86%)
Do not re-use fillers to show witnesses with different suspects *(n = 70)*	55 (79%)
Artificially conceal unusual features (e.g. tattoos, scarring, etc.) for consistency [a] *(n = 69)*	31 (45%)
Artificially re-create unusual features (as above) for consistency [a] *(n = 68)*	13 (19%)

[a] So there is no difference between the suspect's and fillers' characteristics

Questions Related Specifically to Showups

Maximum Time Permitted After Crime to Conduct a Showup. Agencies were asked if there is a maximum time limit after a crime is committed during which they are permitted to conduct a showup. Of 75 respondents, just over a third of agencies (37%) reported there was a time limit. Only 5% said there was no time limit, leaving 59% for which this was not specified.

For the agencies with a time limit, 77% reported that it is captured in hours (as opposed to days) and that it averaged about two hours (ranging from 45 minutes to eight hours).

Location of Showup Procedures. Across all agency types and sizes, the most common practice and location for administering showups is for the witness to be transported to the detained person/suspect who is to be viewed from the backseat of the police vehicle, while the least common is bringing the witness to a police facility to view the detained person/suspect (see **Table 25**). It is important to note that these were not mutually exclusive, so an agency may allow any number of these procedures.

Table 25. Commonality of various procedures for conducting showups

Procedure	Never done	Uncommon	Somewhat or Very Common
Witness may be transported to the detained person/suspect and viewed from the backseat of the police vehicle *(n = 74)*	7 (9%)	11 (15%)	55 (74%)
Witness may view the detained person/suspect at the crime scene *(n = 73)*	4 (5.5%)	31 (42%)	38 (52%)
Detained person/suspect may be transported to the witness for viewing *(n = 73)*	31 (42%)	18 (25%)	24 (33%)
Witness may be brought to a police facility to view the detained person/suspect *(n = 74)*	42 (57%)	23 (31%)	9 (12%)[a]

[a]For this item, no respondents replied "very common"

Handcuffs Allowable in Showups. Across all agency types and sizes, over two-thirds (67%) of the 73 responding agencies reported allowing showups to be done when the identified person or suspect is in handcuffs. For 8 of 10 responding state agencies, this was allowable, as compared to just 64% of local police departments. Too few sheriffs' agencies responded to this question to provide meaningful results.

Witness Description of Perpetrator Required. Agencies were asked if they require that witnesses describe the perpetrator before viewing a detained person/suspect in a showup and 74 responded. Among them, 74% said they did, whereas 26% reported not having this requirement.

Challenges Faced by LEAs in Implementing Policy/Practice Changes

Showups. To assess the challenges faced by agencies over the past decade, agencies were asked to describe any and explain how they overcame them. A total of 48 agencies included comments on the challenges they faced, although a handful did not address specific challenges. A few themes that emerged from the responses received were: (1) the need for re-training of officers, (2) technology, specifically outfitting officers with body worn cameras for documentation and recording of the procedures, (3) challenges associated with documentation consistent with state standards, (4) the need for new data collection forms, and (5) legal issues (e.g., temporary detention, new case law).

A few specific responses representing some of the themes and other unique responses follow:

- "Policy updated to best practice standards."
- "The new practice required significant training in all regards" (cited academy, in service, and training bulletins).
- "A challenge for our agency was ensuring all sworn personnel fill out the proper forms, completing the showup procedure correctly and documenting the procedure properly."
- We were required to shorten the window of time for when a showup can be conducted, which led us to rely more on photo arrays. (*paraphrased*)

Photo Arrays. The same question was asked about challenges faced in the past decade for photo arrays. Fully 57 agencies (35%) wrote in unique answers to this question. There were many agencies that had comments that spanned more than one theme. The key themes identified were (1) uncertainties about the changes, opinions, standards, and laws; (2) the need for re-training; (3) difficulty in finding an independent administrator for blinded or double-blind procedures; (4) challenges faced with adopting sequential procedures like fewer identifications; and (5) technology challenges.

Some of the representative responses are presented below:

- "Challenges in understanding what is a 'best practice.'"
- "The constant changes and opinions about witness identification."
- "Keeping up with ever-changing case law."
- Our state mandated new policies and procedures as well as training, but no one has created the training. (*paraphrased*)
- "Ensuring that everyone received formal and informal training."
- When we switched from sequential presentation, accurate identifications dropped to 10%; now we rarely use either approach. (*paraphrased*)
- Since our agency switched from the 6-pack array lineup to sequential, there have been fewer positive identifications. However we have stronger investigative tools and procedures, so we don't rely solely on eyewitness procedures. (*paraphrased*)
- "It is a bit more cumbersome to ensure the photo array is administered by someone blind to the investigation."
- "Conducted train the trainer classes, then conducted smaller classes with investigators."
- "[Have had to] go back and forth with computer selected photos depending on software available."
- Social media concerns.

Discussion

Over at least the past 25 years, highly regarded organizations and scholars in the U.S. have reviewed the scientific evidence and made recommendations to LEAs regarding policies and procedures. In science and practice more generally, the term "best practices" has become widely used, despite its various definitions and interpretations. Some of the agencies we surveyed used the term "best practices" in their comments throughout, yet few cited the source for those best practices or which specific practices they considered best.

According to Veselý (2011), the term best practice recommendation (BPR) was conceived in the management literature in the early 1990s. However, there are inconsistent definitions and interpretations of the term best practice recommendation and varied approaches to identifying them. Although often used interchangeably with evidence-based policy and/or practices (EBP), BPR is a field-based determination using organizational exemplars to define

what works best, instead of practices informed by science (presumably the most recent scientific consensus). Two of the primary limitations of defining a best practice per Veselý (2011) are the lack of theory and the fact that what is a best practice may be context dependent (what works in one place may not work in another).

Eyewitness researchers from Canada promoted using BPRs because they encourage uniform procedures within a country and evidence-based identification procedures (Bertrand et al., 2018). Yet, one assertion in their article is that the sequential presentation method is a BPR, even though not one of the most respected organizations[61] issuing guidance, standards, or recommendations touts a superior method. Instead, the NRC stated that:

> In recent decades, many police agencies and prosecutors have adopted sequential presentation of photographs, based on the ***belief*** [emphasis added] that this approach improves the performance of an eyewitness. Currently, however, there is no consensus among law enforcement professionals as to whether the sequential presentation procedure is superior to the simultaneous procedure.
> (NRC, 2014, p. 24).

In our view, these beliefs were borne out of earlier research and subsequent advocacy that sequential methods were superior (Steblay, et al., 2003), findings that were not unwarranted at the time. However, since then, the sequential superiority finding has been questioned by scholars who have relied upon improved approaches and further examined the factors that play into the assessment of diagnostic superiority. The present scientific consensus is that there is no sequential advantage, and if anything, there may be a slight simultaneous presentation advantage related to accuracy.

Despite considerable inconsistency in what makes up a "best practice," there are

[61] NIJ (1999), NRC (2014), IACP (2016), CALEA (2023), and APLS (2020)

generally accepted, evidence-based recommendations, some of which have stood the test of time, and others have emerged over time. As such, we aimed to examine current eyewitness practices and policies related to photo arrays and showups in the U.S. since the last nationwide survey was conducted by PERF in 2013, and to assess the extent to which they reflect current science.

One area of consensus is that agencies should adopt formalized, written policies for eyewitness identification procedures. This led to one of the most important findings of our survey: most LEAs now have written policies for both photo arrays and showups. This finding stands in contrast to the very low levels of formal policy adoption reported previously by PERF. Another notable finding is that since the 2013 survey, a slightly lower proportion of agencies have reported using photo arrays. However, the proportion of agencies implementing showups has remained virtually unchanged during this same period.

Earlier in this report, we pointed out that some experts discourage the use of showups, largely due to their assertions that showups are 'inherently suggestive.' Indeed, one agency respondent indicated that showups were too suggestive and not a best practice. Yet, little research has tested this assumption, and undoubtedly, the field has not heeded those warnings to date, as the procedures have continued at about the same rate since 2013. The likely reason is that showups may be the only option in cases where there is little other evidence upon which to act or further investigate. This does not mean that agencies should rely solely upon the result of a showup, as it appears that the majority of prosecutors require corroborating evidence.

Nevertheless, despite continued claims about the suggestive nature of showups, various procedures can reduce the suggestibility of any eyewitness practice (NIJ, 1999). However, the assertion that those administering showups believe they have the correct suspect may not fully account for the practical realities of how showups are conducted in the field. That assertion

assumes that the officer performing the procedure has already formulated an opinion of the likely guilt of the person matching the description, but that appears overstated. In practice, *"be on the lookout"* (BOLO) alerts are typically issued when a witness provides a sufficiently detailed description of the suspect. Generally, the officers or deputies patrolling the area are not the detectives or investigators in the criminal case. [62] Accordingly, they often have not been provided with context other than the description of a potential suspect and the availability of a witness to view a person that matches the perpetrator's description. The accompanying suggestion that witnesses must believe the police have "the" perpetrator has not been adequately tested in the field, again underscoring the overstatement about the "inherent" suggestive nature of showups. It is recommended that as agencies continue to rely on showups, they also work to reduce suggestiveness as they should in all identification procedures. This can be achieved by implementing prior recommendations, such as ensuring the witness does not view the suspect in handcuffs, providing instructions that the investigation will continue regardless of identification, informing the witness that the person they will view may or may not be the perpetrator, and emphasizing that clearing an innocent person is just as important as identifying a guilty individual.

Responding agencies provided illuminating information regarding the source of guidance they rely upon to develop and/or modify eyewitness practices. The finding that agencies greatly rely on prosecutors, professional associations, and state recommendations suggests a need to target dissemination efforts to these groups. Because these organizations are essential in shaping LEA agency policies, the scientific community should continue to focus on improving the

[62] There was one response to an open-ended question indicating that the agency waits for the case detective to conduct the showup, but this is uncommon per numerous anecdotal accounts from agencies.

dissemination and translation of research into relevant policy and practice formation, using clear and concise findings tailored to these specific policy audiences. While reasonable, the reliance by some on case law suggests the need for scientists to provide guidance more broadly to these audiences[63] perhaps through their professional associations like the American Bar Association, the American Judicature Society, the American Judges' Association, and others.

While it is clear from our results that there are many respected sources of guidance on eyewitness policies and practices, this can also lead to confusion as to which guidance to follow, especially due to the lack of uniform national standards and inconsistent state requirements and recommendations. While NIJ, as part of the U.S. Department of Justice, has been the primary authoritative source of this guidance for a long time, their recommendations and guidelines have not been updated in over a quarter of a century. Our findings suggest that some agencies have continued to rely on their 1999 Guidelines despite considerable new science.

Nevertheless, there is some similarity in "best practice" recommendations by several well-respected groups and organizations and for the same general time period in which the guidance was issued. In **Table 26**, we have summarized some of the key areas of guidance and debate across these sources.

[63] This may be difficult to accomplish given the aforementioned perceived utility of research and scientific evidence in judicial practice.

Table 26. Guidelines, recommendations, and conclusions over the past 25 years

	1999	2023	2014	2016	2020
Eyewitness Practice	**NIJ**	**CALEA**	**NRC**	**IACP**	**APLS**
1. Establish policies/procedures for eyewitness ID	Y	Y	Y	Y	Y
2. Provide standardized instructions to witnesses	Y	Y	Y	Y	Y
3. Minimize bias in photographic lineups (arrays)	Y	Y	Y	Y	Y
3a. Include only one suspect per lineup	Y	Y		Y	Y
3b. Standardize photographs of the lineup members	Y	Y		Y	Y
3c. Use double blind or blinded method	R	Y	Y	Y	Y
4. Document witness statements	Y	Y	Y	Y	Y
4a. Audio-record	Y	Y		Y	
4b. Video-record/Record on body-worn camera	Y	Y	Y	Y	Y
4c. Obtain written statement	Y	Y		Y	
	1999	**2023**	**2014**	**2016**	**2020**
Eyewitness Practice	**NIJ**	**CALEA**	**NRC**	**IACP**	**APLS**
5. Presentation method	R	Y		Y	
5a. Sequential preference	N	Y		N	
5b. Simultaneous preference	N			N	
5c. Allow second viewing in sequential array				N	
6. Ask for confidence	Y	Y		Y	Y
6a. Show-up	Y	Y		Y	Y
6b. Photo array	Y	Y		Y	Y
7. Ask for confidence during the initial eyewitness procedure only, and right after the identification or non-ID					
8. Confidence is related to accuracy					Y
8a. Only for unbiased /fair lineups					Y
8b. Regardless of whether the lineup was fair					N
9. How confidence should be provided	Y	Y	Y	Y	Y
9a. Provided in one's own words	Y	Y	Y	Y	Y
9b. Provided on a numeric/percentage scale					Y
9c. Provided on a verbal scale					Y
10. Avoid or advise against the use of showups	N	N		Y	Y

Yes (Y), No (N), Silent on this matter (), Recommend further research (R)

Returning to some of the key recommendations by NIJ and numerous others over the past 25 years, we discuss the current state of practice for each below.

Develop policies and procedures for eyewitness identification of suspects. Since the PERF survey, we have seen a significant improvement in policy development, with 92% of surveyed agencies now reporting that they have a formal, written policy for photo arrays and showups.

Develop and/or implement standardized instructions to witnesses. In our study, 21% of agencies conducting showups reported providing standardized *written* instructions to witnesses (a slight drop since PERF's finding of 27%), although 56% of those administering photo arrays did so (a moderate improvement since PERF's finding of 40%). However, there has been a substantial improvement regarding the provision of standardized *verbal* instructions, with 73% reporting implementing them for showups and 85% for photo arrays compared to just 39% and 43% as previously reported by PERF (2013).

Record witness statements. In the current survey, 53% of agencies reported requiring videotaping for showups and 48% for photo arrays, as compared to the rates reported in the 2013 study of just 32% and 21%, respectively. Also the current rates for agency mandates are notably higher than the proportion of state mandates for videotaping.

Caution that the person in the eyewitness procedure may or may not be the perpetrator. In both our survey and that of PERF (2013), most agencies provide this caution in their instructions. We found very little difference in these rates over time.

Obtain and document a statement of confidence/certainty. It is unclear why fewer agencies conducting showups request confidence (73%) than did so over a decade ago when PERF conducted their survey (85%). It is possible that more agencies are using body-worn camera footage for the procedures and assume that their use is sufficient without asking a

question about confidence. Nevertheless, the proportions for photo arrays remained largely unchanged over the period.

Train all law enforcement officers in eyewitness identification.[64] Regarding the ongoing recommendation that LEAs provide training on eyewitness identification practices to their sworn officers, we found a significant increase in the proportion of agencies doing so since PERF conducted its survey a decade ago. Specifically, for showups, 85% of agencies reported providing this training as compared to 70% in 2013. For photo arrays, 83% reported providing training to sworn personnel, versus just 68% in 2013, reflecting a 15% increase overall.

Barriers to implementing current evidence-based recommendations include inconsistent guidance from various sources, complex nuances in analytical methods that may hinder practitioners' understanding of the state of the science, and ongoing scientific debate surrounding some issues.

An important factor responsible for the reinterpretations and changes to the 1999 NIJ Guidelines arose due to an increased focus on theory, especially signal detection theory, applied to the eyewitness problem (e.g., Gronlund & Benjamin, 2018; Wixted & Mickes, 2012). Applying signal detection theory to eyewitness evidence has updated and refined older recommendations by highlighting the limitations of the performance metrics upon which the field previously relied. One such performance metric, the diagnosticity ratio, confounded an eyewitness's ability to discriminate a guilty suspect from an innocent suspect with an eyewitness's willingness to positively identify someone from a lineup or showup. Another metric, Spearman's rank correlation, masked the true extent of the confidence-accuracy

[64] Few details about the specific curriculum for eyewitness identification training have been provided other than the NIJ trainer's manual, see the archive version from the Department of Education.

relationship predicted by signal detection theory and revealed by calibration analysis (Juslin et al., 1996; Mickes, 2015). As a result, recommended procedures that encouraged eyewitnesses to be more cautious in making positive identifications were erroneously interpreted as indicators of superior performance (Clark et al., 2014; Gronlund et al., 2012; Wixted & Mickes, 2012).

Relative judgment theory (G. Wells, 1984) was the theoretical framework that governed the field before the introduction of signal detection theory. Relative judgment theory proposed that eyewitnesses tend to make comparisons among lineup members, choosing the individual from a lineup that most resembles their memory of the perpetrator. This is fine if the police have a suspect in the lineup who committed the crime. However, if the police have an innocent suspect, the eyewitness will still make a relative judgment and may choose the individual who most resembles their memory of the perpetrator, even if that resemblance is not very strong. This is because someone in the lineup will always be a 'best match' even if that individual is not the correct one.

The focus of relative judgment theory, coupled with the DNA exonerations that began occurring in the late 1980s to mid-1990s, starting with two pivotal cases—Gary Dotson, whose conviction was vacated[65]/charges dropped in Illinois in 1989, and Ronald Cotton, who was exonerated in North Carolina in 1995 after 10.5 years in prison—shifted the focus towards protecting the innocent. Eyewitness recommendations were designed with this goal in mind. However, signal detection theory provides a means to simultaneously weigh false and correct identifications (identifications of an innocent and a guilty suspect, respectively) in a manner not made clear by relative judgment theory.

[65] In Dotson's case not only was DNA important in proving his innocence, but also the victim's testimony that she fabricated the crime and so the charges were dropped against Dotson, but not until after he had served 8 years in state prison and survived a 12-year ordeal. He did not receive a full pardon until 2003.

Signal detection theory, or what Gronlund and Benjamin (2018) have referred to as the new-science view, facilitates the use of eyewitness evidence to help balance protecting the innocent against implicating the guilty. The current understanding of the eyewitness problem and the updated recommendations emphasize that balance. The issues highlighted by signal detection theory, along with numerous studies debating the superiority of sequential or simultaneous presentation methods, may have prompted the NRC to conduct a thorough review of the field, nearly 15 years after the NIJ recommendations were released.

NPI's eyewitness research project had two primary purposes. The first was to take stock of current law enforcement practices across the United States regarding lineups and showups. Given that the last such survey was conducted a decade ago (PERF, 2013), we sought to determine the extent to which more recent recommendations have been adopted by LEAs around the U.S. and the extent to which the updated recommendations have been adopted.

The second was to examine the relationships between witness confidence, latency of response, and accuracy in showups. We did so by conducting both a lab study of showups and the first-ever field study of showups with a focus on likely guilt or innocence (see Amendola & Slipka, 2009; Amendola et al., 2014, for how we estimated ground truth in the field study).

We look first at the factor that helped spur the updated recommendations and the adoption of a new theoretical framework for understanding the eyewitness problem. Then we turn our attention to those updated recommendations and an examination of how extensively they have been adopted by U.S. law enforcement.

As reported herein, the NIJ report, *Eyewitness Evidence: A Guide for Law Enforcement* (1999) was based on the combined efforts of a working group of leading researchers and practitioners. The guidance from that work promoted eyewitness reliability and accuracy, given

the scientific knowledge at the time. Of course, a lot has changed in the almost 25 years since the original recommendations. Many recommendations continue to be supported, while others have challenged prior knowledge and are still being debated and researched. Some enduring practices include developing policies and procedures for eyewitness identification, providing standardized instructions, ensuring only one suspect per lineup, and recording witness statements.

The most recent guidance was provided in 2020 by a subcommittee appointed by the Executive Committee of the American Psychology-Law Society (Division 41 of the American Psychological Association). The subcommittee updated the recommendations, building on the findings of the 2014 NRC report (G. Wells et al., 2020). Four recommendations were maintained from 1998, and five new recommendations were added. The four that were maintained involved the selection of fillers (individuals in a lineup known to be innocent of the crime), pre-lineup instructions, double-blind administration, and collecting a confidence statement. Next, we briefly review the rationale provided for the four maintained recommendations, followed by a review of the five new recommendations made by the APLS.

APLS Maintained Recommendations

Filler Selection. There is a strong consensus that a lineup should be fair (e.g., Colloff et al., 2016). A fair lineup consists, in part, of the selection of fillers that match the suspect to some degree. Current research continues to investigate how similar these matches should be to the suspect (e.g., Carlson et al., 2019; Colloff et al., 2021; McKinley & Peterson, 2023; Shen et al., 2024), and how the fillers should be selected. Regarding the latter, fillers could be selected to resemble the suspect (suspect-matched, Tunnicliff & Clark, 2000) or the perpetrator's description (description matched, Wells et al., 1993). Clark et al. (2013) found no evidence that description-matched filler selection was better. They recommended a combination of both, which is what

many police departments appear to do (Wise et al., 2011). Colloff et al. (2021) put forth an interesting proposal recently. They tested a prediction derived from signal detection theory and showed that selecting fillers dissimilar in appearance to the suspect from a pool of description-matched photos enhanced discriminability. In sum, there is widespread agreement that the fillers should match the suspect, but research continues on the optimal degree of match and how fillers should be chosen to achieve that degree of match.

Pre-Lineup Instructions. The police should inform an eyewitness that the perpetrator may or may not be in the lineup. The meta-analysis by Clark et al. (2014) showed that these "unbiased" instructions induce a conservative shift in the willingness to select from the lineup (i.e., a conservative shift in response bias) but yielded no improvement in discriminability (accuracy). Mickes et al. (2017) evaluated directly the effects of different lineup instructions and compared the outcome of those instructions to identifications made at different confidence ratings. In two of their instruction conditions, standard biased ("If you see the person...please pick him; otherwise, choose the 'not present' option") or unbiased ("The person...may or may not be in the lineup. If you see the person...please pick him; otherwise, choose the 'not present' option"), did not differ from one another in discriminability, consistent with the conclusions of Clark et al. (2014). As predicted by signal detection theory, instructional manipulations and variations in response confidence show similar relationships, with no evidence that unbiased instructions lead to superior performance. Given these results, it might seem puzzling that the NRC Report recommended the use of unbiased instructions. However, the recommendation does make sense from the perspective of balancing costs (implicating the guilty) and benefits (the protection of the innocent). This is especially the case if the cost of convicting an innocent person is greater than the cost of failing to convict a guilty one (see Clark et al., 2015).

Double-Blind Administration. Double-blind testing (e.g., Charman & Quiroz, 2016; Kovera & Evelo, 2020) reduces pressure to choose that might otherwise be exerted by a lineup administrator. This has two potential benefits: (1) It emboldens eyewitnesses to adopt a more conservative response criterion (i.e., only choosing when they are more confident).(2) Double-blind testing also limits the ability of a lineup administrator to steer an eyewitness (implicitly or explicitly) toward selecting a particular individual (Clark et al., 2013, 2015). Although Clark et al. (2013) showed that the absence of double-blind testing can enhance accuracy in some circumstances (it is easier to steer witnesses toward a guilty suspect), sanctioning steering is contrary to the independence of various forms of evidence (e.g., Hasel & Kassin, 2009) as well as principles of procedural justice (see discussion by Clark, 2012). Eyewitness evidence arising from a lineup in which a lineup administrator influences an eyewitness's selection does not deliver an independent contribution to subsequent determinations of a defendant's guilt or innocence.

Collection of a Confidence Statement. The 1998 reforms (G. Wells et al., 1998) recommended that an immediate confidence statement be collected, but this was mainly recommended to safeguard against confidence inflation due to feedback (e.g., G. Wells & Bradfield, 1998). At that time, confidence, even uninflated, was thought to be only a modest predictor of accuracy. This conclusion arose due to the reliance on a measure that masked the degree of the confidence-accuracy relationship (see Juslin et al., 1996). However, the view that confidence and accuracy are, at best, modestly related has been questioned in the last 15 years (e.g., Brewer & G. Wells, 2006; for a review see Wixted & Wells, 2017). Confidence can be a surprisingly reliable predictor of witness accuracy, if the criminal justice system's focus remains

on the initial, immediate, unbiased test conducted on the eyewitness (Wixted et al., 2015; Wixted & Mickes, 2022).

Neither the NIJ Guide, the NRC Report, nor the APLS recommendations specify a preference for simultaneous (viewing all lineup members at once) versus sequential (viewing lineup members one at a time) lineups. However, recent research supports a consensus building around the superiority of simultaneous lineups from the lab (Meisters et al., 2018); Wixted et al., 2015) and the field (Amendola & Wixted, 2015a, b; but see Wells et al., 2015). This is the case despite laboratory data supporting no discriminability difference between the two presentation methods (e.g. Kaesler et al., 2020, 2024).

APLS New Recommendations

Pre-Lineup Interview. One of the new recommendations from G. Wells et al. (2020) calls for law enforcement to conduct a pre-lineup witness interview and document the witness's reports, ideally by video-recording the entire interview. Many scientists have recommended best practices for conducting this interview (e.g., Fisher et al., 2014) and things to avoid (e.g., leading questions, see Loftus, 2018).

The interview should establish an evidence-based suspicion by presenting clear, articulable evidence that reasonably infers a particular person likely committed the crime. This serves as justification for subjecting an individual to identification procedures. The reason for the recommendation is to reduce the frequency of innocent suspects being exposed to identification procedures, thereby increasing the likelihood that guilty suspects are correctly identified.

Video Recording. All identification procedures should be video recorded to document the conduct of the identification procedure, including pre-lineup instructions and the witness's confidence statement. Kassin (1998) discussed two reasons for this recommendation. First, it

provides a means to supplement incomplete (or inaccurate) police reports or recollections of what was done. Second, it ensures best practices have been followed. Other benefits of having a video record of the identification include aiding judges in reviewing motions to suppress an identification, as well as benefits to the prosecution (in evaluating the strength of a case) and the defense (advising a client to accept a plea). Having a video recording also aids jurors' evaluations of witness accuracy (Reardon & Fisher, 2011), despite the fact that very few cases go before juries.

Repeated Identification Attempts. Testing a suspect multiple times enhances the witness's memory for the chosen individual, irrespective of whether that individual committed the crime. Therefore, repeated identification attempts must be avoided (Wixted et al., 2021). There is only one chance to evaluate a witness's memory, and subsequent attempts only contaminate the original memory. At least two related processes can contribute to an initial identification attempt contaminating subsequent identification attempts. One involves a source monitoring error (Johnson et al., 1993) or what Loftus (1976) called unconscious transference. Familiarity with an individual is enhanced by repeated exposure (e.g., the false fame effect, Jacoby et al., 1989; or the bystander effect, Ross et al., 1994) and exacerbated when coupled with the poor memory for the source of that familiarity (e.g., is this the individual who robbed me or this the person I found on Facebook who my friend thinks robbed me?). A second process is the commitment effect. If an individual from a mug book (Goodsell et al., 2009) or an initial lineup (Deffenbacher et al., 2006), is repeated in a subsequent identification procedure, witnesses will tend to stay committed to their initial choice because they are unable to make a second decision that is independent of the initial one.

Reliability. The final new recommendation of the APLS concerns the purported suggestive nature of showups and their reliability:

> Showups should be avoided whenever it is possible to conduct a lineup (e.g., if probable cause exists to arrest the person, then a showup should not be conducted). Cases in which it is necessary to conduct a showup should use the procedural safeguards that are recommended for lineups, including the elimination of suggestive cues, a warning that the detained person might not be the culprit, video-recording the procedure, and securing a confidence statement. (G. Wells et al., 2020, pp. 8-9)

A showup involves the presentation of a single suspect (not accompanied by fillers) to the eyewitness. As discussed earlier in this report, showups have been criticized for their inherently suggestive nature because, as some argue, eyewitnesses obviously know that the person presented is the police suspect (Goodsell et al., 2013; Kassin et al., 1989; Steblay et al., 2003, Wagenaar & Veefkind, 1992; Yarmey et al., 1994). For example, G. Wells et al. (2020) assume that the unreliability of showups is a given, writing, "showups are *clearly* less reliable than lineups…" (p. 27, emphasis added).

However, the evidence for the suggestive nature of showups is lacking, and this may stem from a misunderstanding of how they are conducted in the real world absent field studies examining them. A showup is an easy way to test memory soon after a crime has occurred (thereby quickly apprehending a guilty suspect or releasing an innocent suspect without further delay). In fact, G. Wells et al. (2020) have not disagreed with the latter, and have described ways to reduce showup suggestiveness. For example, a pre-showup instruction—offering additional opportunities to view another individual if they do not identify the detained individual—may reduce mistaken identifications with little reduction to accurate identifications (Eisen et al., 2017; Smith et al., 2018).

The APLS Guidelines indicate that a lineup should be conducted instead of a showup "whenever it is possible," (G. Wells et al., 2020), however, constructing fair lineups is complex. Showups are often more time-sensitive and dependent on specific circumstances. Will the lineup an officer constructs in the heat-of-the-moment be fair? A study by Key et al. (2017) suggests that it might be better to conduct a showup than a biased lineup. They had subjects view a mock crime video, make a timed identification decision from a fair or biased lineup or showup, and then report a confidence judgment. Witnesses in biased lineups were fast and confident, regardless of accuracy. Notably, however, the quickest and most confident witnesses made equally accurate identifications decisions from fair lineups and showups, and both were better than the decisions from biased lineups. No one recommends using biased lineups, but if sufficient lineup fairness is not achieved in the field, it might be better to conduct a showup.

Key et al. (2017) found in a laboratory study that witness confidence collected immediately after a showup identification may help distinguish correct from false identifications, just as it does for lineups. In another ecologically valid showup simulation, a pilot study by Seale-Carlisle (2017) found that high confidence showup identifications were highly accurate (97%). For a contrary finding see Eisen et al. (2022). This accuracy rate is higher than in laboratory studies involving photographs of suspects, where high confidence identifications were 82% accurate in Mickes (2015), 75% accurate in Wetmore et al. (2015), and 85% accurate in Key et al. (2017). Several possible reasons for this discrepancy include the use of live actors which increases fidelity over photographs.

The new science of eyewitness memory offers a different perspective on the faulty identification of Ronald Cotton. In Jennifer Thompson's initial photo lineup, she reported "*I think* this is the guy," only raising her confidence level as the detective pressed her for greater

certainty. The district attorney requested a second live lineup test, in which Cotton was the only person present in both procedures. Again, Thompson took several minutes before stating that Cotton "looks the most like him," indicating that she was highly confident only after the detective asked if she was certain. The jury heard Jennifer declare in court that she was "absolutely sure" in her identification instead of hearing Jennifer's initial report (from Garrett, 2011b). This phenomenon of becoming more confident over time has been displayed in many cases.

This is not how Jennifer Thompson's identification attempt would likely unfold today. According to the new science of eyewitness memory and the recommendations we have reviewed, the police would test her memory only once, in a double-blind manner, accompanied by instructions indicating that the perpetrator may or may not be present. She would make a confidence statement immediately after the identification, which is the confidence statement upon which the criminal justice system would rely. Moreover, if Jennifer's confidence level failed to reach a pre-determined level (e.g., at least moderately or perhaps 80 - 90% confident), which it almost certainly would have failed to achieve, it would have been treated as a non-identification. Consequently, Ronald Cotton might never have been pursued by the prosecution or convicted, and Bobby Poole, the actual perpetrator, might have been apprehended before he terrorized additional women.[66] The old science view of eyewitness memory reinforces the wrong message regarding Jennifer Thompson's identification. Rather than concluding that eyewitnesses can never be trusted, it was by way of her slow, unconfident identifications of Cotton that Jennifer signaled that there was a good chance that she was making an error, something that

[66] Some reports indicate that in the nine months after Thompson's rape, while Ronald Cotton was in prison, Poole committed over 20 more crimes, including multiple rapes and burglaries.

police at the time were not aware of, as eyewitness research was in its infancy. Consequently, false convictions like the one that ensnared Ronald Cotton can be decreased by properly handling eyewitness evidence and a better understanding of eyewitness memory.

Expected Applicability of the Research

The survey results and our report from the lab and field studies (Amendola, et al., 2025) associated with this NIJ grant should provide helpful information detailing the complexity of the available policy and practice guidance on eyewitness identification in criminal investigations. It also should help LEAs (whether local, state police, investigative bodies, or county sheriffs) identify the most common practices in similarly situated agencies, and how those have changed over time as they work to adopt the most effective policies and practices for photo arrays and field showups.

Also, the results of this research may spark greater collaboration among scholars and practitioners, consistent with the NRC committee's keen note that *"The field would benefit from collaborative research among scientists and law enforcement personnel in the identification and validation of new best practices that can improve eyewitness identification procedures" (p. 2).*

The initial working group of the National Institute of Justice (1999) established a baseline for translation of science to practice. It has served as the premier authoritative source for guidance due to its federal status, connection with federal, state, and local law enforcement, and significant means for dissemination efforts, not to mention their ability to bring together a team of highly respected scholars, leading law enforcement professionals, and others with considerable information to bear on the issue. A similar, updated effort, whether led by NIJ or another organization revered by the law enforcement community, is recommended to bring

together a diverse array of organizations and individuals to establish updated and integrated guidance across respected organizations upon which LEAs rely.

This recommendation is not to minimize the numerous available authoritative sources, but to come up with universally accepted and informed practices that would be heeded by the courts, law enforcement organizations, and policy-making organizations that develop standards. These individuals and groups could include legal and psychological scholars, federal, state, and local law enforcement leadership, including prosecutors, judges, police chiefs, sheriffs, and state commissioners, defense attorneys, state legislators or criminal justice commissioners, along with professional associations issuing policy guidance, accreditation bodies, and particularly trainers, especially state Peace Officer Standards Training (POST) leaders (via the International Association of Directors of Law Enforcement Standards and Training, IADLEST).

Training is critical to ensuring evidence-based practices, state regulations, and local policies are followed. In 2003, NIJ published "*Eyewitness Evidence: A Trainer's Manual for Law Enforcement*" (NIJ Training Manual, 2003) to assist law enforcement trainers. It defined and specified minimum performance levels for each procedure and provided the materials needed to establish and maintain the knowledge and skills for conducting the procedures recommended in the 1999 Guide. However, the narrative surrounding eyewitness identification has changed significantly over the last two decades. Although training personnel on eyewitness practices continues to be a consistent recommendation across guidelines, it is largely unknown what guidance is relied upon to design an evidence-based curriculum and/or implement training. New research in several areas has increased understanding of the eyewitness problem. Yet, it is not apparent how much of that guidance has reached LEAs or especially training directors either at training academies, for in-service training, or within investigative units. Perhaps more

importantly, there appears to be no current, standardized training module for eyewitness identification, which can lead to highly inconsistent training across agencies, regions, and states. This, too, should be addressed by any new collaborative effort.

Limitations

As with any study, this report is not without its limitations, nor was the survey itself. While this NPI survey was modeled after the PERF study funded by NIJ and published in 2013, not all PERF's original questions were included in this new study. Part of the reason was so that the NPI research team could pursue additional areas of particular interest, elaborate on some questions, and minimize survey length despite new questions. Specifically, the NPI research team aimed to understand which sources of guidance law enforcement agencies relied upon for policy adoption and modification, while considering a range of field-based recommendations from multiple authoritative sources. Of interest was also the evolving science over the past 15 years regarding presentation methods (sequential vs. simultaneous) used in photo arrays and the more recent findings associated with the relationship between witness confidence and accuracy. As a result of this limitation, comparing findings in this NPI survey with the findings from the 2013 PERF survey was not always possible.

While our response rate was about half that attained by PERF over a decade ago, we have provided context around the diminished response rates and current science that suggests that the impacts of non-response bias are not as significant as once believed, especially when a sample is as large as ours was. Nevertheless, we advise caution in interpreting or generalizing findings for comparisons of groups when a respondent group consists of fewer than 15 agencies.

Another limitation was our inability to compare agencies in this study over time. As we were replicating PERF's stratified, random sampling process, some agencies that had previously

responded to the original PERF survey, were not identified in the new randomization results. Nevertheless, we were given a list of the 619 respondents to the original survey and sent the same survey to them. However, the results of that respondent group were not the subject of this report due to budgetary and time constraints. We hope to provide a panel comparison of those agencies that responded both times in a subsequent report or article.

While we were aware of the confusion over the meaning of double-blind, blinded, and non-blind presentation methods and their conflation with presentation methods, we limited our questions in these areas and accordingly, we were not able to fully evaluate the prevalence of each practice. Also, given the length of our survey and the less predominant practices, we excluded questions about mugshot searches, live lineups, and composite sketches in favor of more detailed questions about photo arrays and showups (the two most common practices).

Finally, drawing conclusions about the influence various sources have on policy change was not possible due to the numerous guidelines and recommendations upon which agencies have relied, and the inability to isolate which source(s) might be responsible for an agency's adoption of certain policies or practices or whether an agency consulted a combination of sources. We did, however, ask what sources agencies consulted, and the answers provide a wealth of information that can inform future dissemination efforts.

The opinions, findings, and conclusions are those of the authors and do not represent the position of the National Institute of Justice, the National Policing Institute, or any of the authors' employing organizations. Questions should be directed to Dr. Karen L. Amendola, National Policing Institute.

References

Albright, T. D. (2023). A scientist's take on scientific evidence in the courtroom. *Proceedings of the National Academy of Sciences*, *120*(41), e2301839120. https://doi.org/10.1073/pnas.2301839120

Albright, T. D., & Garrett, B. L. (2022). The law and science of eyewitness evidence. *Boston University Law Review, 102*, 511. https://www.bu.edu/bulawreview/files/2022/03/ALBRIGHT-GARRETT.pdf

Amendola, K.L., Mickes L., Valdovinos-Olson, M., Carlson, C., Gronlund, S., & Chen, X. (2025). *Confidence, latency, and accuracy in eyewitness identifications made from showups: Evidence from a laboratory experiment and a field study.* Research Report. Washington DC: National Institute of Justice. Office of Justice Programs, U.S. Department of Justice.

Amendola, K. L., & Slipka, M. G. (2009). Strength of evidence scale. *Unpublished instrument. Police Foundation, Washington, DC*. https://www.policinginstitute.org/wp-content/uploads/2020/04/About-the-Strength-of-Evidence-Scale.pdf

Amendola, K. L., & Wixted, J. T. (2015a). Comparing the diagnostic accuracy of suspect identifications made by actual eyewitnesses from simultaneous and sequential lineups in a randomized field trial. *Journal of Experimental Criminology*, *11*, 263-284. https://doi.org/10.1007/s11292-014-9219-2

Amendola, K. L., & Wixted, J.T. (2015b). No possibility of a selection bias, but direct evidence of a simultaneous superiority effect: A reply to Wells et al.. *Journal of Experimental Criminology, 11*, 291–294). https://doi.org/10.1007/s11292-015-9227-x

American Psychology-Law Society, see Wells et al. (2020)

Arizona State University, College of Law. (1974). *Model rules: Eyewitness identification*. Washington, DC: Police Foundation. https://www.ojp.gov/ncjrs/virtual-library/abstracts/model-rules-eyewitness-identification-project-law-enforcement

Arndorfer, A., & Charman, S. D. (2022). Assessing the effect of eyewitness identification confidence assessment method on the confidence-accuracy relationship. Psychology, Public Policy, and Law, 28(3), 414–432. https://doi.org/10.1037/law0000348

Ash, S., Nix, B., & Steinberg, B. (2022). Report on nonresponse bias during the COVID-19 period for the Consumer Expenditures Interview Survey. *Published as part of the Consumer Expenditure Surveys Program Report Series*. https://www.bls.gov/cex/research_papers/pdf/ash-nix-steinberg-covid-nonresponse-diary-report.pdf

Bertrand, M. I., Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Kalmet, N., & Melsom, E. I. (2018). Examining how lineup practices of Canadian and U.S. police officers adhere to their national best practice recommendations. Man. LJ, 41, 1. https://themanitobalawjournal.com/wp-content/uploads/articles/MLJ_41.3/Examining%20How%20Lineup%20Practices%20of%20Canadian%20and%20U.S.%20Police%20Officers%20Adhere%20to%20Their%20National%20Best%20Practice%20Recommendations.pdf

Blomberg, T. G. (2021, December 1). Translational Criminology and Politics. Presentation for

The Justice Research Statistics Association Virtual Training, December [PowerPoint slides]. Florida State University. https://criminology.fsu.edu/sites/g/files/upcbnu3076/files/JRSA%20Blomberg%202021_11-30-21.pdf

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*(1), 11. https://doi.org/10.1037/1076-898X.12.1.11

Brooks, N. (1983). *Police Guidelines: Pretrial Eyewitness Identification Procedures*. Law Reform Commission of Canada. https://www.ojp.gov/pdffiles1/Digitization/93479NCJRS.pdf

Carlson, C. A., Lockamyeir, R. F., Carlson, M. A., Goodsell, C. A., Jones, A. R., Wooten, A. R., & Brand, Z. T. (2025). Comparing the strength of the confidence-accuracy versus response time-accuracy relationship for eyewitness identification. *Scientific Reports*, *15*(1), 11064. https://doi.org/10.1038/s41598-025-96224-y

Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science*, *7*(3), 238-259. https://doi.org/10.1177/1745691612439584

Clark, S. E., Brower, G. L., Rosenthal, R., Hicks, J. M., & Moreland, M. B. (2013). Lineup administrator influences on eyewitness identification and eyewitness confidence. *Journal of Applied Research in Memory and Cognition*, *2*(3), 158-165. https://doi.org/10.1016/j.jarmac.2013.06.003

Clark, S. E., Moreland, M. B., & Gronlund, S. D. (2014). Evolution of the empirical and

theoretical foundations of eyewitness identification reform. *Psychonomic Bulletin & Review*, *21*, 251-267. https://doi.org/10.3758/s13423-013-0516-y

Clark, S. E., Benjamin, A. S., Wixted, J. T., Mickes, L., & Gronlund, S. D. (2015). Eyewitness identification and the accuracy of the criminal justice system. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 175-186. https://doi.org/10.1177/2372732215602267

Collins, W. (2024). *The case for eyewitness identification reform.* Chicago: American Bar Association, Criminal Justice Section. ISBN-13: 978-1639055203. https://www.americanbar.org/products/inv/book/444457311/?login

Colloff, M. F., Wilson, B. M., Seale-Carlisle, T. M., & Wixted, J. T. (2021). Optimizing the selection of fillers in police lineups. *Proceedings of the National Academy of Sciences*, *118*(8), e2017292118. https://doi.org/10.1073/pnas.2017292118

Colloff, M. F., Wade, K. A., & Strange, D. (2016). Unfair lineups make witnesses more likely to confuse innocent and guilty suspects. *Psychological Science*, *27*(9), 1227-1239. https://doi.org/10.1177/0956797616655789

Connors, E., Lundregan, T., Miller, N., & McEwen, T. (1996). *Convicted by juries, exonerated by science: Case studies in the use of DNA evidence to establish innocence after trial.* Research Report. NCJ 161258). Washington DC: National Institute of Justice. Office of Justice Programs, U.S. Department of Justice. https://www.ojp.gov/pdffiles/dnaevid.pdf

Czajka, J. L., & Beyler, A. (2016). Background paper declining response rates in federal surveys: Trends and implications. *Mathematica policy research*, *1*, 1–86. https://aspe.hhs.gov/sites/default/files/private/pdf/255531/Decliningresponserates.pdf

Davern, M. (2013). Nonresponse rates are a problematic indicator of nonresponse bias in survey

research. *Health services research*, *48*(3), 905. https://pmc.ncbi.nlm.nih.gov/articles/PMC3681235/pdf/hesr0048-0905.pdf

Deffenbacher, K.A., Bornstein, B.H. & Penrod, S.D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30*, 287–307. https://doi.org/10.1007/s10979-006-9008-1

Dodson, C. S., & Dobolyi, D. G. (2016). Confidence and eyewitness identifications: The cross-race effect, decision time and accuracy. *Applied Cognitive Psychology, 30*(1), 113–125. https://doi.org/10.1002/acp.3178

Eisen, M. L., Smith, A. M., Olaguez, A. P., & Skerritt-Perta, A. S. (2017). An examination of showups conducted by law enforcement using a field-simulation paradigm. *Psychology, Public Policy, and Law*, *23*(1), 1. http://dx.doi.org/10.1037/law0000115

Eisen, M. L., Ying, R. C., Chui, C., & Swaby, M. A. (2022). Comparing witness performance in the field versus the lab: How real-world conditions affect eyewitness decision-making. *Law and Human Behavior*, *46*(3), 175. https://doi.org/10.1037/lhb0000485

Fisher, R. P., Schreiber Compo, N., Rivard, J., & Hirn, D. (2014). Interviewing witnesses. *The SAGE handbook of applied memory*, 559-578. https://sk.sagepub.com/hnbk/edvol/the-sage-handbook-of-applied-memory/chpt/interviewing-witnesses

Fosnacht, K., Sarraf, S., Howe, E., & Peck, L.K. (2017). How Important are High Response Rates for College Surveys? *The Review of Higher Education 40*(2), 245–265. https://doi.org/10.1353/rhe.2017.0003

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong.* Harvard University Press. https://doi.org/10.4159/harvard.9780674060982

Goodsell, C. A., Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2013). Showups. In B. L. Cutler (Ed.), *Reform of eyewitness identification procedures* (pp. 45–63). American Psychological Association. https://doi.org/10.1037/14094-003

Goodsell, C. A., Neuschatz, J. S., & Gronlund, S. D. (2009). Effects of mugshot commitment on lineup performance in young and older adults. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(6), 788-803. https://doi.org/10.1002/acp.1512

Green, L. W., Ottoson, J. M., García, C., Hiatt, R. A., & Roditis, M. L. (2014). Diffusion theory and knowledge dissemination, utilization and integration. *Frontiers in public health services & systems research*, *3*(1), 3. https://pmc.ncbi.nlm.nih.gov/articles/PMC4524659/pdf/nihms683033.pdf

Green, L. W., Ottoson, J. M., Garcia, C., & Hiatt, R. A. (2009). Diffusion theory and knowledge dissemination, utilization, and integration in public health. *Annual Review of Public Health*, *30*(1), 151–174. https://doi.org/10.1146/annurev.publhealth.031308.100049

Gronlund, S. D., & Benjamin, A. S. (2018). The new science of eyewitness memory. In *Psychology of learning and motivation* (Vol. 69, pp. 241-284). Academic Press. https://doi.org/10.1016/bs.plm.2018.09.006

Gronlund, S. D., Carlson, C. A., Dailey, S. B., & Goodsell, C. A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology: Applied*, *15*(2), 140. https://doi.org/10.1037/a0015082

Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., & Graham, M. (2012). Showups versus lineups: An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition*, *1*(4), 221-228. https://doi.org/10.1016/j.jarmac.2012.09.003

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using receiver operating characteristic analysis. *Current Directions in Psychological Science*, *23*(1), 3-10. https://doi.org/10.1177/0963721413498891

Gross, S. R. & Shaffer, M. (2012). Exonerations in the U.S., 1989 – 2012. National Registry of Exonerations. https://www.law.umich.edu/special/exoneration/documents/exonerations_us_1989_2012_full_report.pdf

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public opinion quarterly*, *72*(2), 167-189. https://doi.org/10.1093/poq/nfn011

Hendra, R., & Hill, A. (2019). Rethinking response rates: new evidence of little relationship between survey response rates and nonresponse bias. Evaluation review, 43(5), 307–330. https://doi.org/10.1177/0193841X18807719

Hinojosa, Cintia, & Nesterak, Evan (2021). Science. Special Collection: Examining the Intersection of Behavioral Science and Advocacy. *Behavioral Scientist.* https://behavioralscientist.org/special-collection-examining-the-intersection-of-behavioral-science-and-advocacy/

Innocence Project of New Orleans (n.d.). Eyewitness Misidentification. https://ip-no.org/criminal-justice-issues/wrongful-convictions/

International Association of Chiefs of Police (2016). *Eyewitness Identification: Model Policy, Concepts & Issues Paper, and Need to Know*. https://www.theiacp.org/sites/default/files/2018-08/EyewitnessIDPolicy2016.pdf

IACP (2020). Eyewitness Identification 2016 Model Policy and Concepts and Issues Discussion, Summary. Alexandria, VA: IACP. https://www.theiacp.org/iacp-policy-center-topic-directory

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of personality and social psychology*, *56*(3), 326.

Johnson, M. T. (1993). Memory phenomena in the law. *Applied cognitive psychology*, *7*(7), 603-618. https://doi.org/10.1002/acp.2350070705

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1304. https://psycnet.apa.org/buy/1996-05780-017

Kaesler, M., Dunn, J. C., Ransom, K., & Semmler, C. (2020). Do sequential lineups impair underlying discriminability? *Cognitive research: principles and implications*, *5*, 1-21. https://doi.org/10.1186/s41235-020-00234-5

Kassin, S. M. (1998). Eyewitness identification procedures: The fifth rule. *Law and Human Behavior. 22*(6), 649–653. https://doi.org/10.1023/A:1025702722645

Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The" general acceptance" of

psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, *44*(8), 1089. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0003-066X.44.8.1089

Kaesler, M., Dunn, J. C., & Semmler, C. (2024). Clarifying the effects of sequential item presentation in the police lineup task. *Cognition*, *250*, 105840. https://doi.org/10.1016/j.cognition.2024.105840

Krieger, N., LeBlanc, M., Waterman, P. D., Reisner, S. L., Testa, C., & Chen, J. T. (2023). Decreasing survey response rates in the time of COVID-19: implications for analyses of population health and health inequities. *American Journal of Public Health*, *113*(6), 667-670. https://doi.org/10.2105/AJPH.2023.307267

Küfner, B. Sakshaug, J. W., & Zins, S. (2022, December). Analysing establishment survey non-response using administrative data and machine learning. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *185*(2), S310–342. https://doi.org/10.1111/rssa.12942

Kukucka, J., & Faigman, D. (2024). Shaken baby syndrome has been discredited. Why is Robert Roberson still on Death Row? *Scientific American*. (October 26). https://www.scientificamerican.com/article/shaken-baby-syndrome-has-been-discredited-why-is-robert-roberson-still-on/

Lasker, E. (1954). Possible Procedural Safeguards against Mistaken Identification by Eyewitnesses. *UCLA Law Review*, 2, 552.

Lindsay, R. C., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of applied psychology*, *70*(3), 556. https://psycnet.apa.org/buy/1985-30824-001

Loftus, E. F. (2018). Eyewitness science and the legal system. *Annual Review of Law and Social Science*, *14*(1), 1-10. https://doi.org/10.1146/annurev-lawsocsci-101317-030850

Loftus, E. F. (1976). Unconscious transference in eyewitness identification. *Law & Psychol. Rev.*, *2*, 93. https://www.ojp.gov/ncjrs/virtual-library/abstracts/unconscious-transference-eyewitness-identification

Lombrozo, T. (2017). What makes science science? National Public Radio (NPR). [Radio broadcast] January 30. In 13.7, Cosmos & Culture, Commentary on Science and Society. https://www.npr.org/sections/13.7/2017/01/30/512402110/what-makes-science-science#:~:text=Far%20from%20undercutting%20the%20value,scientific%20knowledge%20is%20continually%20evolving

Lum, C., Maupin, C., & Stoltz, M. (2023). The supply and demand shifts in policing at the start of the pandemic: A national multi-wave survey of the impacts of COVID-19 on American law enforcement. *Police quarterly*, *26*(4), 495-519. https://doi.org/10.1177/10986111221148217

Macias, W., Springston, J. K., Weaver, R. A. L., & Neustifter, B. (2008). A 13-year content analysis of survey methodology in communication related journals. *Journal of current issues & research in advertising*, *30*(1), 79-94. https://doi.org/10.1080/10641734.2008.10505240

Manson v. Braithwaite, 432 U.S. 98 (1977). https://supreme.justia.com/cases/federal/us/432/98/

McKinley, G. L., & Peterson, D. J. (2023). Using objective measures to examine the effect of suspect-filler similarity on eyewitness identification performance-Final Registered Report. *Cognitive Research: Principles and Implications*, *8*(1), 68. 023) 8:68 https://doi.org/10.1186/s41235-023-00522-w

McNabb, N. S., Farrell, B. R., & Brown, C. R. (2017). Voluntary Adoption of Evidence-Based Practices by Local Law Enforcement: Eyewitness Identification Procedures in Arkansas, Iowa, Kansas, Missouri, and Nebraska. *J. Gender Race & Just.*, *20*, 509–541 U Iowa Legal Studies Research Paper No. 2017-30

Meisters, J., Diedenhofen, B., & Musch, J. (2018). Eyewitness identification in simultaneous and sequential lineups: An investigation of position effects using receiver operating characteristics. *Memory*, *26*(9), 1297-1309. https://doi.org/10.1080/09658211.2018.1464581

Mickes, L. (2015). Receiver operating characteristic analysis and confidence–accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*(2), 93–102. https://doi.org/10.1016/j.jarmac.2015.01.003

Mickes, L., Clark, S. E., & Gronlund, S. D. (2017). Distilling the confidence-accuracy message: A comment on Wixted and Wells (2017). *Psychological Science in the Public Interest*, *18*(1), 6–9. https://journals.sagepub.com/doi/pdf/10.1177/1529100617699240

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18*(4), 361–376. https://doi.org/10.1037/a0030609

Mickes, L., & Wixted, J. T. (2025). Eyewitness Memory. *Oxford handbook of human memory, Oxford University Press, 2294-2328.* https://doi.org/10.1093/oxfordhb/9780190917982.013.80

Moore, K. N., Pennekamp, P., Yu, C., & Zwemer, D. U. (2024). Eyewitness Confidence Does

Not Necessarily Indicate Identification Accuracy. *Policy Insights from the Behavioral and Brain Sciences*, *11*(2), 149-155. https://doi.org/10.1177/23727322241268384

Morris, Z. S., Wooding, S., & Grant, J. (2011). The answer is 17 years, what is the question: understanding time lags in translational research. *Journal of the Royal Society of Medicine*, *104*(12), 510–520. https://journals.sagepub.com/doi/pdf/10.1258/jrsm.2011.110180

National Public Safety Information Bureau. *Directory of Law Enforcement Administrators* (2020). https://www.safetysource.com/directories/

National Institute of Justice (1999). *Eyewitness Evidence: A Guide for Law Enforcement*. U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Technical Working Group for Eyewitness Evidence. NCJ 178240. https://www.ojp.gov/pdffiles1/nij/178240.pdf

National Research Council. (2014). *Identifying the Culprit: Assessing Eyewitness Identification*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18891 https://www.nap.edu/catalog/18891/identifying-the-culprit-assessing-eyewitness-identification

Neil v. Biggers, 409 U.S. 188 (1972). https://supreme.justia.com/cases/federal/us/409/188/

Nielson, K. R., Zhang, Y., & Ingram, J. R. (2022). The impact of COVID-19 on police officer activities. *Journal of Criminal Justice*, *82*, 101943. https://doi.org/10.1016/j.jcrimjus.2022.101943

Nix, J., Pickett, J. T., Baek, H., & Alpert, G. P. (2019). Police research, officer surveys, and response rates. *Policing and Society*, *29*(5), 530–550.

https://digitalcommons.unomaha.edu/cgi/viewcontent.cgi?article=1064&context=criminaljusticefacpub

Norman, D. A. (2010). The research-practice gap: The need for translational developers. *Interactions*, *17*(4), 9–12. https://dl.acm.org/doi/10.1145/1806491.1806494

NRC, see National Research Council.

PBS (2018). FRONTLINE: What Jennifer Saw. [Radio broadcast]. Documentary. https://www.pbs.org/wgbh/pages/frontline/shows/dna/interviews/thompson.html

PERF, see Police Executive Research Forum,

Police Executive Research Forum (2013). *A National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies.* Washington, D.C.: PERF. https://www.ojp.gov/pdffiles1/nij/grants/242617.pdf

Quigley-McBride, A., & Wells, G. L. (2023). Eyewitness confidence and decision time reflect identification accuracy in actual police lineups. *Law and Human Behavior*, *47*(2), 333. https://psycnet.apa.org/manuscript/2023-44292-001.pdf

Reardon, M. C., & Fisher, R. P. (2011). Effect of viewing the interview and identification process on juror perceptions of eyewitness accuracy. *Applied Cognitive Psychology*, *25*(1), 68-77. https://doi.org/10.1002/acp.1643

Riley, Patricia A. (2019). *An Overview of the Research on Eyewitness Identification of Strangers, New and Old.* Prosecutors' Center for Excellence. See: https://pceinc.org/wp-content/uploads/2019/12/20191210-Overview-of-Research-on-EWID-Evidence-Pat-Riley.pdf

Riley, Patricia A. (2018). *Analysis of Scientific Research on Eyewitness Identification*. U.S. Attorney's Office, District of Columbia. Prosecutor's Center for Excellence. See:

https://pceinc.org/wp-content/uploads/2019/11/20180100-ANALYSIS-OF-SCIENTIFIC-RESEARCH-ON-EWID-AND-OPPOSITION-TO-HENDERSON-INSTRUCTION-DC-US-Attorneys-Office-AUSA-Patricia-Riley.pdf

Ross, D. F., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and lineup identification: Toward a memory blending approach. *Adult eyewitness testimony: Current trends and developments*, 80-100.

Rouzer, S. K., Kalinowski, L. M., & Kaseda, E. T. (2023). The importance of promoting scientific advocacy & outreach for trainees. *Neuropsychopharmacology*, *48*(5), 713–715. https://doi.org/10.1038/s41386-023-01530-6

Rubin, R. (2023). It takes an average of 17 years for evidence to change practice—the burgeoning field of implementation science seeks to speed things up. *JAMA*, *329*(16), 1333–1336. https://doi.org/10.1001/jama.2023.4387

Sauer, J. D., Palmer, M. A., & Brewer, N. (2019). Pitfalls in using eyewitness confidence to diagnose the accuracy of an individual identification decision. *Psychology, Public Policy, and Law*, *25*(3), 147. https://doi.org/10.1037/law0000203

Seale-Carlisle, T. M. (2017). *A Re-examination of Eyewitness Memory Phenomena Using Receiver Operating Characteristic Analysis and Confidence-Accuracy Characteristic Analysis* (Doctoral dissertation, Royal Holloway, University of London).

Seale-Carlisle, T. M., Colloff, M. F., Flowe, H. D., Wells, W., Wixted, J. T., & Mickes, L. (2019). Confidence and response time as indicators of eyewitness identification accuracy in the lab and in the real world. *Journal of Applied Research in Memory and Cognition*, *8*(4), 420–428. https://doi.org/10.1016/j.jarmac.2019.09.003

Semmler, C., Dunn, J., Mickes, L., & Wixted, J. T. (2018). The role of estimator variables in

eyewitness identification. *Journal of Experimental Psychology: Applied*, *24*(3), 400. https://doi.org/10.1037/xap0000157

Shen, K. J., Huang, J., Lam, A. L., & Wixted, J. T. (2024). The effects of filler similarity and lineup size on eyewitness identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.https://doi.org/10.1037/xlm0001342

Shih, T. H., & Fan, X. (2008). Comparing response rates from web and mail surveys: A meta-analysis. *Field methods*, *20*(3), 249-271. https://doi.org/10.1177/1525822X08317085

Smith, A. M., Wells, G. L., Lindsay, R. C. L., & Myerson, T. (2018). Eyewitness identification performance on showups improves with an additional-opportunities instruction: Evidence for present–absent criteria discrepancy. *Law and Human Behavior, 42*(3), 215–226. https://doi.org/10.1037/lhb0000284

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *25*, 459-473. https://link.springer.com/article/10.1023/a:1012888715007

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law*, *17*(1), 99. https://doi.org/10.1037/a0021650

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 27*(5), 523–540. https://doi.org/10.1023/A:1025438223608

Stoddard, G. D. (1935). The lag between practice and research. *The Phi Delta Kappan*, *17*(6), 180–82. https://www.jstor.org/stable/20258410

Thompson-Cannino, J., Cotton, R., & Torneo, E. (2010). *Picking Cotton: Our Memoir of Injustice and Redemption*. St. Martin's Press.

Tunnicliff, J. L., & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions?. *Law and Human Behavior*, *24*(2), 231-258. https://doi.org/10.1023/A:1005463020252

U. S. Department of Justice (2017). Eyewitness Identification: Procedures for Conducting Photo Arrays. (Yates Memo) https://www.justice.gov/archives/opa/press-release/file/923201/dl?inline

Veselý, A. (2011). Theory and methodology of best practice research: a critical review of the current state. *Central European Journal of public policy*, *5*(02), 98-117

Wagenaar, W. A., & Veefkind, N. (1992). Comparison of one-person and many-person lineups: A warning against unsafe practices. *Psychology and Law: International Perspectives*, 275-285. https://doi.org/10.1515/9783110879773

Wåhlberg, A. E. A., & Poom, L. (2015). An empirical test of nonresponse bias in internet surveys. *Basic and Applied Social Psychology*, *37*(6), 336–347. https://doi.org/10.1080/01973533.2015.1111212

Weber, N., Brewer, N., Wells, G. L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: the unruly 10-12-second rule. *Journal of Experimental Psychology: Applied*, *10*(3), 139. https://doi.org/10.1037/1076-898X.10.3.139

Weir, K. (2016, February 1). Mistaken identity. *Monitor on Psychology*, *47*(2). https://www.apa.org/monitor/2016/02/mistaken-identity

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator

variables. *Journal of Personality and Social Psychology, 36*(12), 1546. https://doi.org/10.1037/0022-3514.36.12.1546

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*(2), 89–103. https://doi.org/10.1111/j.1559-1816.1984.tb02223.x

Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior, 44*(1), 3–36. https://doi.org/10.1037/lhb0000359

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, *78*(5), 835. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0021-9010.78.5.835

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, *22*(6), 603. https://doi.org/10.1023/A:1025750605807

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Key, K. N., & Goodsell, C. A. (2015). Do the clothes make the criminal? The influence of clothing match on identification accuracy in showups. *Journal of Applied Research in Memory and Cognition*, *4*(1), 36-42. https://doi.org/10.1016/j.jarmac.2014.12.003Get rights and content

Wise, R. A., Sartori, G., Magnussen, S., & Safer, M. A. (2014). An examination of the causes and solutions to eyewitness error. *Frontiers in Psychiatry*, *5*, 102. https://doi.org/10.3389/fpsyt.2014.00102

Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives on*

Psychological Science, *7*(3), 275-278. https://doi.org/10.1177/1745691612442906

Wixted, J. T., & Mickes, L. (2022). Eyewitness memory is reliable, but the criminal justice system is not. *Memory*, *30*(1), 67-72. https://doi.org/10.1080/09658211.2021.1974485

Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger III, H. L. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist*, *70*(6), 515.

Wixted, J. T., Mickes, L., Dunn, J. C., Clark, S. E., & Wells, W. (2016). Estimating the reliability of eyewitness identifications from police lineups. *Proceedings of the National Academy of Sciences*, *113*(2), 304-309. https://www.pnas.org/doi/full/10.1073/pnas.1516814112

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. https://journals.sagepub.com/doi/full/10.1177/1529100616686966

Wixted, J. T., Wells, G. L., Loftus, E. F., & Garrett, B. L. (2021). Test a Witness's Memory of a Suspect Only Once. *Psychological Science in the Public Interest*, 22(1_suppl), 1S-18S. https://doi.org/10.1177/15291006211026259

Wu, M. J., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports*, *7*, 100206. https://doi.org/10.1016/j.chbr.2022.100206

Yarmey, A. D., Yarmey, A. L., & Yarmey, M. J. (1994). Face and voice identifications in showups and lineups. *Applied Cognitive Psychology*, *8*(5), 453-464. https://doi.org/10.1002/acp.2350080504

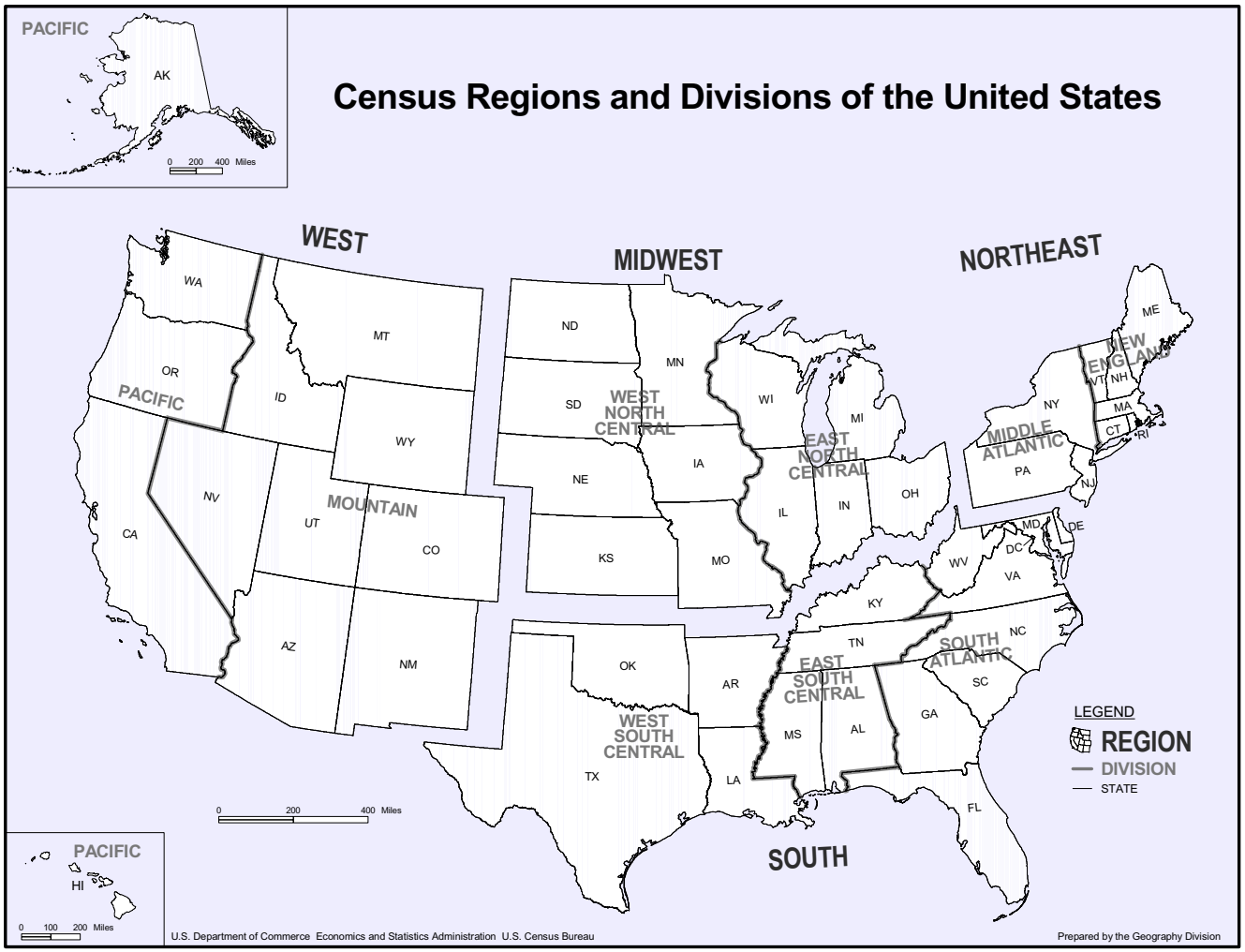
Yates, Sally Q. (2017). See U.S. Department of Justice.

Appendices

Appendix A: NDLEA Strata Sizes with Selected Sample Count

Region	Department Type	Department Size	Population Count	Sample Count
West	Local Police	1 to 25	717	20
		26 to 50	164	20
		51 to 99	125	20
		100 to 499	101	20
		500 or more	5	5
		Missing	13	13
	County Sheriffs	1 to 25	205	20
		26 to 50	73	20
		51 to 99	27	27
		100 to 499	36	20
		500 or more	7	7
		Missing	3	3
South	Local Police	1 to 25	3,235	20
		26 to 50	485	20
		51 to 99	254	20
		100 to 499	185	20
		500 or more	12	12
		Missing	30	20
	County Sheriffs	1 to 25	756	20
		26 to 50	273	20
		51 to 99	154	20
		100 to 499	188	20
		500 or more	16	16
		Missing	1	1
Northeast	Local Police	1 to 25	1,825	20
		26 to 50	424	20
		51 to 99	160	20
		100 to 499	103	20
		500 or more	1	1
		Missing	19	19
	County Sheriffs	1 to 25	60	20
		26 to 50	31	20
		51 to 99	16	16
		100 to 499	21	21
		500 or more	3	3
		Missing	1	1
Midwest	Local Police	1 to 25	2,875	20
		26 to 50	407	20
		51 to 99	151	20
		100 to 499	69	20
		500 or more	2	2
		Missing	37	20
	County Sheriffs	1 to 25	689	20
		26 to 50	123	20
		51 to 99	59	20
		100 to 499	47	20
		500 or more	2	2
		TOTAL	**14,190**	**769**

Appendix B: States within Regions defined by the U.S. Census Bureau



Appendix C: Weighting Strategy

In our design, we included several distinct sets of agencies:

1. State police, highway patrol, and state investigative bureaus
2. Other municipal law enforcement agencies and county sheriffs

Some agencies in the first set also appear in the second set. Additionally, some agencies in the first set are no longer active and/or have reorganized (e.g., by consolidating with other agencies).

To create the weights, we took three steps:

1. Compute design weights
2. Compute non-response weights
3. Combine design and non-response probabilities for a total weight

All state police, highway patrol, and investigative bureaus were included, so no sampling took place for these agencies. Ultimately, there were 74 of these agencies, and 31 responded.

All other agencies were sampled based on a stratification scheme similar to the one used by PERF in their original sampling. The agencies were drawn from the National Directory of Law Enforcement Agencies (NDLEA). The total population of municipal law enforcement agencies, both local police (including town, city, and county police), and county sheriff offices, is 14,379. Of this population, 765 agencies were sampled, and 156 agencies responded.

The table below shows the stratification scheme, with the total population size, the number of units included in the sampling frame, and the number of responses we received.

Region	Department Type	Department Size	Population Count	Sample Count	Random Respondent
West	Municipal Law Enforcement	1-25	719	20	6
West	Municipal Law Enforcement	26-50	164	20	7
West	Municipal Law Enforcement	51-99	127	20	7
West	Municipal Law Enforcement	100-499	102	21	6
West	Municipal Law Enforcement	500	5	5	4
West	Municipal Law Enforcement	Unknown	13	13	0
West	County Sheriffs	1-25	218	20	4
West	County Sheriffs	26-50	76	20	4
West	County Sheriffs	51-99	37	27	1
West	County Sheriffs	100-499	51	19	5
West	County Sheriffs	500	9	7	2
West	County Sheriffs	Unknown	28	3	0
South	Municipal Law Enforcement	1-25	2926	15	2
South	Municipal Law Enforcement	26-50	441	18	3
South	Municipal Law Enforcement	51-99	228	16	6
South	Municipal Law Enforcement	100-499	186	17	9
South	Municipal Law Enforcement	500	17	16	6
South	Municipal Law Enforcement	Unknown	29	19	1
South	County Sheriffs	1-25	722	17	0
South	County Sheriffs	26-50	254	15	0
South	County Sheriffs	51-99	151	18	3
South	County Sheriffs	100-499	190	17	4
South	County Sheriffs	500	16	9	2

South	County Sheriffs	Unknown	1	1	0
Northeast	Municipal Law Enforcement	1-25	1828	20	2
Northeast	Municipal Law Enforcement	26-50	425	20	3
Northeast	Municipal Law Enforcement	51-99	160	20	3
Northeast	Municipal Law Enforcement	100-499	105	20	9
Northeast	Municipal Law Enforcement	500	1	1	0
Northeast	Municipal Law Enforcement	Unknown	19	19	0
Northeast	County Sheriffs	1-25	76	20	2
Northeast	County Sheriffs	26-50	33	19	3
Northeast	County Sheriffs	51-99	20	16	2
Northeast	County Sheriffs	100-499	33	21	2
Northeast	County Sheriffs	500	7	3	1
Northeast	County Sheriffs	Unknown	2	1	0
Midwest	Municipal Law Enforcement	1-25	3193	24	6
Midwest	Municipal Law Enforcement	26-50	456	23	4
Midwest	Municipal Law Enforcement	51-99	178	24	9
Midwest	Municipal Law Enforcement	100-499	82	23	8
Midwest	Municipal Law Enforcement	500	2	2	0
Midwest	Municipal Law Enforcement	Unknown	42	21	1
Midwest	County Sheriffs	1-25	732	23	3
Midwest	County Sheriffs	26-50	145	24	4
Midwest	County Sheriffs	51-99	68	22	7
Midwest	County Sheriffs	100-499	57	22	5
Midwest	County Sheriffs	500	5	4	0

The design weights were based on the probability that a unit was included in the sampling frame. The weights were calculated as the inverse of this probability. For instance, there were 719 municipal law enforcement agencies in the West with between 1 and 25 officers. Our sampling frame had 20 such agencies. Thus, an agency in this stratum had a little less than a 3% chance of being included in the sampling frame (20/719 = 0.0278). The design weight for these units is, therefore, 35.95 (1/0.0278), meaning that each unit in our sampling frame had a corresponding weight of about 36 respondents in the population.

In the next stage, we computed non-response weights. These weights accounted for the fact that some agencies were more likely to respond to the survey than others. When we received responses from agencies that were less likely to respond, the weights allowed us to place more value on those respondents. To estimate this probability, we used logistic regression with predictors for department size, department region, and department type (with some agencies given the designation 'Unknown' for these categories for which we lacked information). Estimated non-response probabilities ranged from .037 (County Sheriffs in the West of 50-100 sworn officers) to .8 (Municipal Law Enforcement Agencies in the West of 500 or more sworn officers). Again, the response weights were the inverse of these probabilities.

Finally, to compute the total weight, the design and response probabilities were combined through multiplication to yield an overall probability of having a response. Inverting this probability yields a total weight. We then scaled these weights such that the sum of the weights is equal to the total number of responses in the survey. In the final data frame, the variable design_nr_weights represented the weight scaled to the population (i.e., how many agencies in the population did this response represent), and combined_weights_scaled represented the weight scaled to the sample, which is typical for survey weights.

The other new variables included in the files were the intermediate steps described above:

- design_weights = the design-based weights
- nr_weights= the non-response weights

We also included the variables used when stratifying and computing these probabilities:

region, dept_size_bin_label, and depttype.

These variables were combined in the group to indicate the group to which an agency belonged.

Appendix D: Map of Agency Survey Responses by Region (n = 187)

Distributed 839 total surveys