# a clearer view

# THE IMPACT OF THE NATIONAL ACADEMY OF SCIENCES REPORT ON EYEWITNESS IDENTIFICATION

*BY THOMAS D. ALBRIGHT AND JED S. RAKOFF*

**S**ix years ago, the U.S. National Academy of Sciences (NAS) convened a panel of experts to consider the problem of eyewitness identification. Eyewitnesses have long played a significant role in criminal investigations and prosecutions. Despite this history of valued testimony, mounting evidence from both science[1] and conviction records of DNA-based exonerees[2] indicates that eyewitnesses easily and unwittingly identify innocent suspects.

The societal problem caused by eyewitness misidentification — the conviction of innocent people — stems from a panoply of dysfunction, ranging from scientific naiveté and investigative bias to prosecutorial disregard and judicial ignorance, as well as a natural human tendency to trust what people say they saw.[3] The NAS eyewitness panel was thus composed of experts representing a variety of fields, including the scientific study of human visual perception and memory, sociology, statistics, law, and law enforcement. After reviewing evidence from many sources, the panel released a comprehensive report in the fall of 2014,[4] which identified factors that commonly lead to erroneous conviction and made substantial recommendations for reform.

The recommendations of the NAS eyewitness panel covered three topic areas: scientific understanding; law enforcement practice; and use of eyewitness evidence in the courtroom. We are pleased to report that these recommendations have been embraced by both scientific and legal communities and have had, in a few short years, a significant positive impact on the field. The NAS report itself, which is freely available from the National Academies Press website,[5] has been downloaded more than 13,000 times and has elicited rich discussion, new scientific research, and improved legal practice. In what follows we summarize some of these developments.

## Engagement of the Science Community

The overarching scientific development stimulated by the NAS report has been the engagement of a research community with expertise in sensory and cognitive processes. This development addresses a longstanding weakness, which is that reputed fact and operational strategies in the field of eyewitness identification have often come from applied studies[6] that have little grounding in a principled mechanistic understanding of how people see, remember, and make decisions.

One product of this engagement is a recasting of the problem.[7] Traditionally, factors that affect eyewitness performance have been identified symptomatically and classified as "estimator" or "system" variables, which are taxonomic distinctions based on time of influence and the degree to which the criminal justice system has control over the outcome.[8] Estimator variables characterize the viewing conditions and perceptual/cognitive state of the witness at the time of the crime. These variables (e.g., lighting, viewing distance, stress, and fear) should be considered when assessing the validity of testimony, but they cannot be changed. System variables (e.g., the manner in which a lineup is conducted), by contrast, may influence identification accuracy after the crime and can thus be controlled to improve the likelihood of correct identification. Although insights from this classification scheme have indeed led to improvements, it has promoted a palliative approach to eyewitness performance — we can simply recognize and/or tweak the state of these variables to get the best outcome — at the expense of understanding and mitigating the root causes of eyewitness failure.[9]

## An Information-Processing Approach to Eyewitness Identification

Accurate identification requires that the witness correctly perceive and ▶

remember the events of a crime. Although this task is uncommon for most people, we regularly engage in analogous behaviors that require recognition of a previously seen and remembered target, such as finding the luggage on the carousel, the tumor in the lung, or the car in the parking lot. In all such cases, the observer is an instrument for information measurement, classification, and storage. The root causes of misidentification are thus tied to the operating characteristics of human observers — sensitivity, storage capacity, and susceptibility to interference or bias — which have been studied extensively by the science community for decades. The product of this research — a principled mechanistic understanding of how vision and memory work — suggests ways to optimize human performance of visual recognition and avoid conditions in which people are likely to fail.[10]

At the most basic level, we know that three factors bear on the performance of an eyewitness: uncertainty, bias, and confidence. Uncertainty results from "noise," or unpredictable perturbation of otherwise meaningful signals. Noise is ever-present in our sensory and mnemonic worlds and constrains the information that can be acquired by an observer. Such constraints naturally reduce the accuracy and utility of observations and thus, to the extent that uncertainty can be quantified, place useful upper bounds on the probability that a witness's identification is correct. Usually unbeknownst to the observer, bias quietly fills informational gaps left by uncertainty — we see what we expect to see and are none the wiser. Confidence is the degree to which the identifier feels certain that his observation is correct; its most common manifestation in this context is overconfidence, or a special form of bias in

*AT THE MOST BASIC LEVEL, WE KNOW THAT THREE FACTORS BEAR ON THE PERFORMANCE OF AN EYEWITNESS: UNCERTAINTY, BIAS, AND CONFIDENCE.*

which the observer implicitly rates the certainty of the experience to be greater than it warrants. This is commonly the result of external forces, such as other evidence or opinions,[11] which drive the observer's certainty in line with a larger story.[12] Overconfidence may be the most pernicious problem with eyewitness testimony, since even very poor-quality information can influence decision and action if communicated to others with certainty.[13]

This human information-processing perspective on eyewitness identification had already begun to take shape at the time of the NAS report.[14] One important issue was growing awareness by the eyewitness research community that a lineup identification is the unique product of two unknown variables: the strength of the observer's recognition memory and the criterion used by the observer to decide. Memory strength is affected by uncertainty, and the decision criterion is determined by various biases as well as overconfidence; but, at the end of the day, it can be difficult or impossible to know whether a suspect identification reflects strong memory or a lax decision criterion.[15] This difference matters, since the former is much more likely to yield an accurate identification.

In an effort to overcome the underlying ambiguity of identification decisions, much recent eyewitness research has focused on two specific questions:[16] (1) What can be done to improve the ability of an eyewitness to access his/her memory of events from the crime scene, such that the witness is better able to discriminate the face of the perpetrator from faces of innocent people?; and (2) Is it possible to determine, in the face of manifold forms of uncertainty and bias, the likelihood that any given witness has identified the right person?

### Can We Improve Discriminability of the Culprit?

Most research bearing on this question has been inspired by specific hypotheses about how presentation of a set of choice stimuli — lineup faces in this case — influences memory-based discriminability.[17] Because the question itself is so fundamental, and stimulus presentation is one of the few tractable variables, no topic in recent eyewitness identification research has stirred as much activity, interest, and controversy as the manner in which a lineup is conducted.

Until a few decades ago, lineups were always performed such that all faces were visible at the same time (initially live and later largely with photographs). Laboratory studies were designed to quantify the average performance of observers under these "simultaneous lineup" conditions using a simple metric — termed the "diagnosticity ratio" — defined as the probability of correctly identifying the culprit relative to the probability of incorrectly identifying an innocent suspect.

Motivated by the hypothesis that identification errors result from relative rather than absolute comparisons of lineup faces, eyewitness researchers

designed an alternative to the simultaneous lineup that was intended to curtail relative judgments. This "sequential lineup," in which faces are presented one at a time, was predicted to yield a greater ratio of correct to incorrect identifications.[18] This hypothesis was upheld in early studies, and a number of law enforcement jurisdictions subsequently switched from simultaneous to sequential lineup procedures.[19]

On the surface of things, the move from simultaneous to sequential lineups seemed sensible and appealing, given the relative reduction of misidentifications. The logic motivating the change was deeply flawed, however, as noted in the NAS report, because it failed to consider the fact that a lineup identification confounds the effects of recognition memory and decision criterion. In other words, the cause of the reduction of misidentifications was ambiguous — it was not possible to know whether it had been due to witnesses being better at accessing memory from the crime scene or witnesses simply being more conservative about pointing the finger.[20] As further research showed, sequential lineups encourage witnesses to adopt a more conservative criterion — they simply make fewer identifications of any sort — but there has been no significant evidence that sequential lineups elicit greater discriminability of lineup participants based on recognition memory for the culprit, which is the desired outcome.

### Simultaneous vs. Sequential Lineups: Which Is Better, and Why?

The sequential-simultaneous debate was in full swing when the NAS eyewitness panel was convened. Also by this time, some investigators had adopted a signal detection approach known as ROC analysis[21] to evaluate witness performance in a way that disentangles memory strength from decision criterion. To do so requires holding either memory strength or decision criterion constant at known values. While the criterion that an eyewitness uses to decide may be inscrutable, it is correlated with the witness's confidence in his or her decision.[22] On average, confident witnesses are more likely to be selective and identify only those faces that meet a stringent criterion. It follows that the probability of correct identification, measured for a known set of confidence-based decision criteria, is proportional to the strength of recognition memory.

Getting a handle on one of the two critical variables for identification — the decision criterion — helps to overcome the fundamental ambiguity of the identification process. This approach, which has dominated eyewitness studies of the past few years, has yielded many important insights into the factors that influence accurate identifications.[23] Although this approach yielded tentative support for simultaneous lineups, as studies suggested improved discriminability with this approach, the NAS panel felt that a recommendation in support of either simultaneous or sequential was premature. The panel nonetheless urged caution when considering a change of lineup procedures.

The NAS report prompted a surge of studies designed to evaluate performance as a function of lineup type.[24] Considered together with the earlier simultaneous/sequential comparisons cited in the NAS report, these recent analyses indicate that, on average, witnesses are better able to optimize sensitivity to their memories — that is, they manifest better discriminability — when simultaneous lineups are used.[25]

### Beyond Traditional Lineups?

The relative discriminability afforded by simultaneous or sequential lineups was only made clear by the adoption of a more sophisticated approach to data analysis, but these two basic lineup procedures have been in use for decades. Although both methods are simple to apply in practical settings by law enforcement, and the outcomes are easy to intuit (perhaps deceptively so) by triers of fact, there are good scientific reasons to break outside this box and explore new ways to improve eyewitness performance. For example, the method used to evaluate the merits of traditional lineups measures the ratio of correct-to-false identifications, which is assessed for a known set of confidence-based decision criteria (such as expressed confidence). That method, of course, depends on precise and accurate measurements of witness confidence. As we have seen, confidence can veer off the rails when witnesses are exposed to other sources of information.[26] The alternative would be to develop an approach that nails down the other key variable that underlies lineup identification: strength of recognition memory.

Memory strengths are not directly accessible because they live only within the mind of the observer. But they can be estimated using experimental techniques that have long been part of the repertoire of basic scientific studies of human information processing. These techniques, known as "perceptual scaling," map the relationship between a set of physical stimuli and the corresponding responses of an observer's perceptual system.[27] In the case of eyewitness identification, scaling techniques can be used to quantify perceived similarity of each lineup face to a remembered target. Perceived similarity, in turn, is an estimate of ▶

the recognition memory signal elicited by each face. Because it affords a view beneath the surface of a categorical identification, scaling of recognition memory signals elegantly reduces the identification problem to one of statistical inference,[28] in which lineup faces can be classified probabilistically as perpetrator or innocent suspect based on the estimated memory signals.[29] This new approach holds much promise as a means to overcome the fundamental ambiguity of traditional eyewitness identifications.[30]

## How Can We Tell if a Witness Has Identified the Correct Person?

The aforementioned studies of lineup type shed valuable light on the contributions of recognition memory and decision criteria in lineup identification. In doing so, they pinpoint lineup conditions that yield, on average, the best ability of a witness to discriminate the culprit from innocent suspects. These discoveries are an indispensable basis for policy decisions about the type of lineup to use in actual criminal cases. This approach reveals nothing, however, about the probability that a given witness identification is correct, which is of course what the trier of fact really needs to know.

Probability of correctness, or "accuracy," is distinct from discriminability and is defined as the ratio of correct identifications of the culprit relative to all (correct or incorrect) culprit identifications reported.[31] Accuracy is surely impacted by various forms of uncertainty and bias. For example, if witnesses cannot easily see actors and events of the crime because of dim lighting or distance, they face uncertainty and are less likely to be correct in their identifications. One might suppose that witnesses who are uncertain are less confident, and thus expres-

sions of confidence might usefully predict accuracy. For many years, however, studies revealed little correlation between confidence and accuracy in recognition memory tasks,[32] and legal standards for the use of witness confidence judgments in eyewitness identification naturally followed that scientific foundation.[33]

This seemingly inoperative relationship between confidence and accuracy has been challenged by two recent observations. First, as noted above, confidence is empirically correlated with the decision criterion for identification.[34] Second, as observers become more conservative in their decision criterion — selecting a target only when absolutely certain — they are expected to become more accurate.[35] It follows that confidence should predict accuracy. Indeed, one of the most important discoveries since the release of the NAS report confirms this prediction: On average, highly confident[36] witnesses are in fact highly accurate in their identifications.[37] This is true even in the presence of significant uncertainty and bias, which reduce overall accuracy and correspondingly reduce the overall likelihood that a witness will report a high confidence identification. The bottom line is that a high-confidence identification, when it occurs at the time of the lineup, is likely to be a correct identification, and thus witness confidence is of great probative value for the trier of fact.

Further support for this view comes from a novel procedure for estimating the strength of recognition memory signals based on confidence judgments.[38] In this case, rather than rendering a categorical identification, the witness assigns a confidence rating to each lineup participant, which is presumed to reflect the corresponding strength of recognition memory.

The accuracy of suspect identifications was found to be correlated with confidence in those identifications. The strength of that correlation was greatest when confidence ratings regarding fillers (or lineup participants known to be innocent) were low, and thus not competing with recognition memory for the suspect. In other words, estimates of memory strength for *all* lineup participants further increases the informational value of a high-confidence suspect identification.

Another new approach to the accuracy question, which stems from the perceptual scaling method highlighted above, also seeks to estimate the recognition memory signals that underlie an eyewitness identification. This method involves the presentation of all possible pairs of a set of lineup faces. Witnesses are asked to make relative — not absolute — judgments: Which face of each pair looks more like the perpetrator?[39] The consistency of such judgments across different face-pair presentations serves as an objective quantitative index of certainty for any individual witness. This index is witness-specific and sidesteps the criterion dependence of confidence statements: A witness is asked to choose between two alternatives, rather than offer a more abstract statement about confidence as to an absolute judgment, which is notoriously difficult to quanitfy. Instead, the index of certainty provides a statistical basis for triage of uncertain witnesses and may prove useful for predicting the correctness of identifications.

## Other Key Variables in Eyewitness Research

In addition to the emergence of a broader information-processing perspective and the aforementioned advancements pertaining to eyewit-

ness discriminability and accuracy, there have been a number of more specific scientific developments stimulated by the NAS report. These cover a range of factors that affect eyewitness uncertainty, bias, and confidence.[40] We mention two of these factors here because of their promise for improving eyewitness performance: (1) lineup filler selection and (2) use of other visual cues for lineup identification.

### Filler Selection

Fillers are lineup participants known to be innocent, who serve as lures to challenge recognition memory. The choice of fillers has long been known to markedly influence eyewitness performance.[41] To understand why this is true, it helps to consider eyewitness identification — and object recognition more generally — as a process of statistical inference. People recognize objects probabilistically based on the degree to which they elicit a memory signal corresponding to a particular target previously seen. It naturally follows that similar objects are more likely to elicit the same memory signal and thus have similar likelihoods of being recognized as the target. The similarity of fillers to the suspect is thus an important variable that affects both eyewitness discriminability and accuracy.

Lineups composed of fillers who are all of roughly the same degree of physical or perceived similarity to the suspect are termed "fair." In other words, to be fair, lineup fillers must look like each other as well as like the suspect himself. Conversely, lineups composed of fillers that possess differing degrees of similarity to the suspect are termed "unfair" or "biased." An unfair lineup, in which one filler is closer in similarity to the perpetrator, reduces uncertainty by lessening the number of sensible choices and

*THE ACCURACY OF SUSPECT IDENTIFICATIONS WAS FOUND TO BE CORRELATED WITH CONFIDENCE IN THOSE IDENTIFICATIONS.*

oversimplifies the witness's statistical inference. The eyewitness may be essentially dealing with a two-choice problem rather than a six-choice problem. This has the effect of simultaneously increasing the likelihood of identifying the culprit and the likelihood of misidentifying someone who looks like the culprit — thereby decreasing both discriminability and accuracy.

Despite the well understood and potentially disastrous consequences of unfair lineups, by the time of the NAS report, there had been very few serious attempts to systematize the process of filler selection. Published guidelines for filler selection stated that lineups should be constructed to ensure that "the suspect does not unduly stand out" and should "avoid using fillers that so closely resemble the suspect that a person familiar with the suspect might find it difficult to distinguish the suspect from the fillers."[42] This guidance — fillers should be similar to the suspect but not too much so — is clearly open to interpretation and is often applied by different agents in different ways.

More recently, and for a variety of reasons having nothing specifically to do with eyewitness identification,[43] many studies of face recognition have focused on metrics of face similarity. As applied to eyewitness identification, the goal would be to employ these met-

rics to create lineups in which fillers are all of known similarity to the suspect. There are two kinds of approaches to this problem, one of which uses physical parameters of the face — such as distance between the eyes, height of the forehead and width of the mouth — to define similarity.[44] The other kind of approach defines face similarity perceptually, based on human judgments.

In both approaches, we begin with a library of faces, ideally representative of the demographic of interest. The physical similarity approach exploits the fact that faces commonly differ from one another along multiple physical dimensions. Thus, the characteristics of any face can be quantified and described by a unique point in some high-dimensional "face space." Collapsing this high-dimensional space onto a manageable three dimensions allows one to compute the Euclidean distance between any two points, which serves as a measure of physical similarity between the respective faces.

The alternative perceptual approach measures face similarity directly from reports of human observers. Numerous attempts have been made to do this by asking people to rate face similarity,[45] but these are plagued by the criterion-dependence of subjective ratings and the effort required to apply this method to large face libraries. Perceptual scaling methods, such as the paired comparison procedure described above, avoid the criterion problem and are a natural choice to produce similarity measures for any given pair of faces.

The desired product of both physical and perceptual approaches is a set of similarity measures for all possible pairs of faces in the library.[46] When drawing from this similarity-indexed library, it should be possible ▶

to customize filler selection for a given suspect by specifying both the average face similarity distance between suspect and fillers, and the variance of face similarity amongst the fillers. Such a system could then be used to empirically determine the parameters of lineup face similarity that yield the best eyewitness performance.

### Other Visual Cues for Lineup Identification

Lineups today rarely employ live participants, and with the move to still photographs has come a significant reduction of visual information that might be used for recognition. Lineup photographs are en face, they lack stereoscopic and motion cues that might reveal three-dimensional structure. They are absent whole-body information, such as posture and gait, and they are often monochromatic. At the same time, it has become increasingly clear from studies of visual object recognition that performance is better when more information-bearing cues are available to the observer.[47]

A remarkable recent study along these lines identified the specific pieces of information from facial images that were used by observers to perceptually encode three-dimensional shape and texture.[48] These coding rules were then used to "reverse engineer" new faces, which were found to be perceptually similar to those originally observed. The significance of this for lineup design is that the empirically determined coding principles highlight the types of visual information that would be beneficial for eyewitness identification. Although simplistic studies involving video images for lineup identification have failed to find much utility,[49] a principled approach in which visual presentations of lineup participants convey information that

*THE GOOD NEWS IS THAT IN THE FEW YEARS SINCE THE NAS PANEL ISSUED ITS REPORT, NO FEWER THAN 19 STATES HAVE PASSED LEGISLATION OR HAVE ADOPTED RULES REQUIRING THE REFORMS SET FORTH IN RECOMMENDATIONS . . . . THE BAD NEWS IS THAT 31 STATES STILL HAVE NOT ACTED.*

matches the ways in which people encode and remember faces is likely to be of great value for improving eyewitness performance.

### Beyond Basic Research: Eyewitness Enhancements in the Real World

In addition to kickstarting a lot of new and interesting science that may ultimately improve the ability of eyewitnesses to identify the culprit, the NAS report made recommendations aimed at enhancing and standardizing practices employed by law enforcement and at strengthening the use of eyewitness evidence in the courts. These recommendations have led to specific reforms, which we highlight in the following sections.

### Better Lineup Procedures

As reflected above, improving lineup procedures and, more generally, police techniques regarding eyewitness identifications, can only go so far toward avoiding inaccurate identifi-

cations, which are sometimes caused by factors unrelated to such practices. Nevertheless, misleading police practices are a material factor in a number of misidentifications. Accordingly, the NAS panel made five recommendations to improve police practices, especially in connection with lineups and photo arrays. These recommendations were: (1) training all law enforcement officers on the variables that can affect eyewitness identifications; (2) adopting "blind" lineup and photo array procedures (such as having the procedure administered by an officer who is not involved in the underlying investigation); (3) providing the officers who do administer the procedures with standardized witness instructions designed to avoid suggestiveness and contamination; (4) documenting the witness's stated level of confidence at the time of an identification; and (5) videotaping the witness identification process.

The good news is that in the few years since the NAS panel issued its report, no fewer than 19 states have passed legislation or have adopted rules requiring the reforms set forth in recommendations 2, 3, and 4, i.e., blind procedures, standardized instructions, and recording of confidence levels. These states are California, Colorado, Florida, Georgia, Hawaii, Illinois, Kansas, Louisiana, Maryland, Massachusetts, Montana, Nebraska, Nevada, New Hampshire, New Mexico, New York, Oklahoma, Utah, and West Virginia. Some of these states have also adopted recommendation 5 (videotaping), though others have simply chosen to recommend it where feasible. Finally, while only a few states have adopted recommendation 1 (universal training), implementation of the other reforms has presumably served to sensitize police officers to some of the attendant issues and problems.

In addition to these statewide legislative and regulatory reforms, two criminal justice organizations with broad jurisdiction have recently weighed in on police practices for eyewitness identification. In September 2016, the International Association of Chiefs of Police released a new "Model Policy" for the conduct of lineups.[50] Similarly, in January 2017, the U.S. Department of Justice released new "Procedures for Conducting Photo Arrays,"[51] the first revision of DOJ policies since 1999. In both cases, the procedures adopted were drawn directly from the recommendations of the NAS report.

The bad news is that 31 states still have not acted. While local police authorities in several of these states had already adopted some or all of these best practices even prior to the NAS report, it can only be hoped that the remaining states and municipalities will follow the lead of the states that have now adopted most of the report's recommendations.

### Strengthening Eyewitness Evidence in the Courts

The NAS report also made four recommendations designed to strengthen the value of eyewitness identification evidence in court. These were: (6) making more frequent pre-trial judicial inquiries into the adequacy of the eyewitness testimony proposed to be offered; (7) making juries aware of the circumstances and confidence of the eyewitness's prior identifications; (8) allowing expert witnesses to educate juries about the problems with eyewitness testimony; and (9) alternatively, using jury instructions to convey this information.

Here, the record of improvements has been more spotty. Following the earlier lead of New Jersey, Massachusetts has now issued a set of jury instructions to be given before or after the testimony of an eyewitness to alert juries to some of the potential issues, and the Supreme Court of Utah has now approved a new rule allowing judges to conduct pre-trial suppression hearings to determine if an eyewitness's identification is too problematic to be presented to a jury. But, with these exceptions, the main effect of the NAS report on the courts, thus far, has been to sensitize some (though by no means all) judges to some of the problems.

Nevertheless, anecdotal evidence indicates that, since the issuance of the NAS report, defense counsel have become more assiduous in raising with the courts issues regarding eyewitness identification. For example, in several jurisdictions, appellants have raised as an issue on appeal the failure of trial courts to provide funds for indigent defendants to retain eyewitness experts. While it does not appear that any of these cases has yet provided a definitive answer as to whether such funds should be made available, it may be inferred that issues regarding the problems with eyewitness identifications are at least becoming more salient in the minds of judges.

In short, there has been some progress in the courts, but not an overwhelming response. It is not clear why. After all, the 19 states that have enacted lineup reforms represent a fairly broad cross-section of America, so there does not appear to be an ideological "split" that prevents improvements elsewhere. Perhaps it is just a matter of inertia. But given the stakes involved — i.e., the wrongful conviction of innocent persons — one may well hope for better in the future.

### Coda

The unfortunate products of eyewitness misidentification — frivolous investigations, errant prosecutions, and wrongful convictions — have persisted for many years, largely due to a lack of understanding. The NAS eyewitness report, however, has provided a wealth of information, offering a detailed and coherent roadmap for change by highlighting the associated primary problems and underlying causes. As we have summarized herein, reform is now underway on multiple fronts, fueled by passion, principles, and good ideas. Indeed, this is shaping up to be one of the great success stories at the intersection of science and law. We look forward to the next five years.            ▶

**THOMAS D. ALBRIGHT** is Professor and Conrad T. Prebys Chair at the Salk Institute for Biological Studies in La Jolla, Calif. He served as co-chair of the NAS Committee on Scientific Approaches to Eyewitness Identification and on the National Commission on Forensic Science and is a member of the Human Factors Committee for the National Institute for Standards and Technology (NIST) Organization of Scientific Area Committees for Forensic Science.

**JED S. RAKOFF** is a U.S. District Judge for the Southern District of New York and an adjunct professor at Columbia and NYU Law Schools. He served on the National Commission on Forensic Science and as co-chair of the National Academy of Science's Committee on Eyewitness Identification.

1   *See, e.g.*, Susan T. Fiske, Steven E. Clark, Aaron S. Benjamin, John T. Wixted, Laura Mickes & Scott D. Gronlund, *Eyewitness Identification and the Accuracy of the Criminal Justice System*, 2 Pol'y Insights from Behav. & Brain Sci. 175–86 (2015).

2   *Eyewitness Identification Reform*, Innocence Project, https://www.innocenceproject.org/causes/eyewitness-misidentification (last visited Feb. 21, 2020).

3   *See* Thomas D. Albright, *Why Eyewitnesses Fail*, 114 Proc. Nat'l Acad. Sci. 7758–64 (2017); Brandon Garrett, Convicting the Innocent: Where Criminal Prosecutions Go Wrong (2011).

4   Nat'l Res. Council, Identifying the Culprit: Assessing Eyewitness Identification (2014), https://www.nap.edu/catalog/18891/identifying-the-culprit-assessing-eyewitness-identification.

5   *Id.*

6   Gary L. Wells & Elizabeth A. Olson, *Eyewitness Testimony*, 54 Ann. Rev. Psychol. 277–95 (2003); Gary L. Wells, *Applied Eyewitness-Testimony Research: System Variables and Estimator Variables*, 36 J. Pers. Soc. Psychol. 1546–57 (1978).

7   Albright, *supra* note 3, at 7758–64.

8   Gary L. Wells, *Applied Eyewitness-Testimony Research: System Variables and Estimator Variables*, 36 J. Pers. Soc. Psychol. 1546–57 (1978).

9   By analogy, insulin therapy is a successful long-term palliative approach to Type 1 diabetes, but curing or preventing the disease in the human population will ultimately require a thorough scientific understanding of cell biology and metabolism.

10  Peter H. Lindsay & Donald A. Norman, Human Information Processing: An Introduction to Psychology (1972).

11  For example, "I'm certain that's him. My sister saw the same thing!"

12  Daniel Kahneman, Thinking, Fast and Slow (2011).

13  Eyewitnesses do this with sincerity and without awareness, but it has long been an explicit strategy employed by magicians and autocrats.

14  Laura Mickes, Heather D. Flowe & John T. Wixted, *Receiver Operating Characteristic Analysis of Eyewitness Memory: Comparing the Diagnostic Accuracy of Simultaneous and Sequential Lineups*, 18 J. Experimental Psychol. Applied 361–76 (2012); John T. Wixted & Laura Mickes, *A Signal-Dre-Detection Model of Eyewitness Identification*, 121 Psychol. Rev. 262–76 (2014).

15  Formally this is known as an "ill-posed problem," i.e. a problem for which there is not sufficient information available to derive a unique solution.

16  Laura Mickes, *Receiver Operating Characteristic Analysis and Confidence-Accuracy Characteristic Analysis in Investigations of System Variables and Estimator Variables that Affect Eyewitness Memory*, 4 J. Applied Res. in Memory & Cognition 93–102 (2015).

17  Stronger recognition memory is the causal factor that leads to better discriminability of objects in a recognition task. Wixted & Mickes, *supra* note 14, at 262–76; Andrew P. Yonelinas, Mariam Aly, Wei-Chun Wang & Joshua D. Koen, *Recollection and Familiarity: Examining Controversial Assumptions and New Directions*, 20 Hippocampus 1178–94 (2010); John T. Wixted, *Dual-Process Theory and Signal-Detection Theory of Recognition Memory*, 114 Psychol. Rev. 152–76

(2007); Roderick C.L. Lindsay & Gary L. Wells, *Improving Eyewitness Identifications from Lineups: Simultaneous Versus Sequential Lineup Presentation*, 70 J. Applied Psychol. 556–64 (1985); Gary L. Wells, *The Psychology of Lineup Identifications*, 14 J. Applied Soc. Psychol., 89–103 (1984).

18  Lindsay & Wells, *supra* note 17, at 556–64.

19  Police Executive Research Forum, A National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies (2013).

20  In fact, it is impossible to determine the true cause given only a simple ratio of correct to incorrect identifications.

21  Analysis of Receiver Operating Characteristics (ROC) is a tool initially developed for evaluation of electronic signal communication, in order to describe the performance of a binary signal classifier as a function of the decision criterion. The approach has been employed extensively for evaluation and mechanistic understanding of human sensation and memory. *See* David M. Green & John A. Swets, Signal Detection Theory and Psychophysics (1966); James P. Egan, Recognition Memory and the Operating Characteristic (1958).

22  John T. Wixted & Gary L. Wells, *The Relationship Between Eyewitness Confidence and Identification Accuracy: A New Synthesis*, 18 Psychol. Sci. Pub. Int. 10–65 (2017); Laura Mickes et al., *ROCs in Eyewitness Identification: Instructions Versus Confidence Ratings*, 31 Applied Cognitive Psychol. 467–77 (2017); John T. Wixted, Laura Mickes, Steven E. Clark, Scott D. Gronlund & Henry Roediger, *Initial Eyewitness Confidence Reliably Predicts Eyewitness Identification Accuracy*, 70 Am. Psychol. 515–26 (2015).

23  John T. Wixted & Laura Mickes, *Theoretical vs. Empirical Discriminability: The Application of ROC Methods to Eyewitness Identification*, 3 Cognitive Res.: Principles & Implications 9 (2018); Laura Mickes, *Receiver Operating Characteristic Analysis and Confidence-Accuracy Characteristic Analysis in Investigations of System Variables and Estimator Variables That Affect Eyewitness Memory*, 4 J. Applied Res. in Memory & Cognition 93–102 (2015); John T. Wixted & Laura Mickes, *ROC Analysis Measures Objective Discriminability for Any Eyewitness Identification Procedure*, 4 J. Applied Res. in Memory & Cognition 329–34 (2015); John T. Wixted et al., *supra* note 23, at 515–26; John T. Wixted & Laura Mickes, *A Signal-Detection-Based Diagnostic Feature-Detection Model of Eyewitness Identification*, 121 Psychol. Rev. 262–76 (2014); Curt A. Carlson & Maria A. Carlson, *An Evaluation of Lineup Presentation, Weapon Presence, and a Distinctive Feature Using ROC Analysis*, 3 J. Applied Res. in Memory & Cognition 45–53 (2014); Laura Mickes et al., *supra* note 14, at 361–76.

24  S. Willing, Birk Diedenhofen & Jochen Musch, Presenting a Similar Foil Prior to the Suspect Reduces the Identifiability of the Perpetrator in Sequential Lineups: A ROC-Based Analysis (2019) (unpublished manuscript); Travis M. Seale-Carlisle, Stacy A. Wetmore, Heather D. Flowe & Laura Mickes, *Designing Police Lineups to Maximize Memory Performance*, 25 J. Experimental Psychol.: Applied 410–30 (2019); Julia Meisters, Birk Diedenhofen & Jochen Musch, *Eyewitness Identification in Simultaneous and Sequential Lineups: An Investigation of Position Effects Using Receiver Operating Characteristics*, 26 Memory 1297–1309 (2018);  J. Trent Terrell,

Aaron R. Baggett, Michelle N. Dasse & Karenna F. Malavanti, *Hybridization of Simultaneous and Sequential Lineups Reveals Diagnostic Features of Both Traditional Procedures*, 13 Applied Psychol. in Crim. Just. 97–109 (2017); Travis M. Seale-Carlisle & Laura Mickes, *U.S. Line-ups Outperform U.K. Line-ups*, 3 Royal Soc'y Open Sci. 160300 (2016); Charles A. Goodsell, Ryan M. McAdoo & Robert Lockamyeir, Eyewitness Encoding Strength on Lineup Identification Performance (2015) (unpublished manuscript); Shannon M. Andersen, Curtis A. Carlson, Maria A. Carlson & Scott D. Gronlund, *Individual Differences Predict Eyewitness Identification Performance*, 60 Personality & Individual Differences 36–40 (2014).

25  *See* Seale-Carlisle et al., *supra* note 24, at 410–30 & Fig. 7 (2019).

26  Nancy K. Steblay, Gary L. Wells & Amy B. Douglass, *The Eyewitness Post Identification Feedback Effect 15 Years Later: Theoretical and Policy Implications*, 20 Psychol. Pub. Pol'y & Law 1–18 (2014); Henry L. Roediger, John T. Wixted & K. Andrew DeSoto, *The Curious Complexity Between Confidence and Accuracy in Reports from Memory*, *in* Memory and Law 84–118 (2012).

27  L.L. Thurstone, *A Law of Comparative Judgment*, 34 Psychol. Rev. 273–86 (1927).

28  Michael D. Ward & John S. Alquist, Maximum Likelihood for Social Science: Strategies for Analysis (2018); Wilson S. Geisler, *Ideal Observer Analysis*, *in* The Visual Neurosciences 825–37 (2003); David M. Green & John A. Swets, Signal Detection Theory and Psychophysics (1966).

29  This is analogous, for example, to the use of statistical image classification algorithms to determine the probability that a grainy feature in a satellite photograph is a ballistic missile and not a tree.

30  Sergei Gepshtein & Thomas D. Albright, *Perceptual Scaling Improves Eyewitness Identification*, Psychonomic Society Annual Meeting Abstracts (2019); Sergei Gepshtein, Yurong Wang, Fangcho He, Dinh Diep, Thomas D. Albright, *A Perceptual Scaling Approach to Eyewitness Identification* (under review).

31  By analogy, my accuracy at identifying the hawk flying across my yard is equal to the number of times I correctly identify the hawk relative to all of the times I report seeing a hawk.

32  *See, e.g.*, Siegfried L. Sporer, Steven Penrod, Don Read & Brian Cutler, *Choosing, Confidence, and Accuracy: A Meta-Analysis of the Confidence–Accuracy Relation in Eyewitness Identification Studies*, 118 Psychol. Bulletin 315–27 (1995); Brian R. Clifford & Jane Scott, *Individual and Situational Factors in Eyewitness Testimony*, 63 J. Applied Psychol. 352–59 (1978); Kenneth A. Deffenbacher, Ann L. Brown & William Sturgill, *Some Predictors of Eyewitness Memory Accuracy, in* Practical aspects of memory 219–26 (1978); Michael R. Leippe, Gary L. Wells & Thomas M. Ostrom, *Crime Seriousness as a Determinant of Accuracy in Eyewitness Identification*, 63 J. Applied Psychol. 345–51 (1978).

33  For review, see Wixted & Wells, *supra* note 22, at 10–65.

34  Laura Mickes et al., *ROCs in Eyewitness Identification: Instructions Versus Confidence Ratings*, 31 Applied Cognitive Psychol. 467– 77 (2017); Chad Dubé & Caren M. Rotello, *Binary ROCs in Perception and Recognition Memory Are Curved*, 38 J. Experimental Psychol.: Learning, Memory, & Cogni-

TION 130–51 (2012); Joshua D. Koen & Andrew P. Yonelinas, *From Humans to Rats and Back Again: Bridging the Divide Between Human and Animal Studies of Recognition Memory with Receiver Operating Characteristics*, 18 LEARNING & MEMORY 519–22 (2011); J.P. Egan, Arthur I. Schulman & Gordon Z. Greenberg, *Operating Characteristics Determined By Binary Decisions and By Ratings,* 31 J. ACOUSTICAL SOC'Y OF AM. 768–73 (1959).

[35] This decision criterion-accuracy relationship is an explicit prediction of signal detection theory. *See* DAVID M. GREEN & JOHN A. SWETS, SIGNAL DETECTION THEORY AND PSYCHOPHYSICS (1966).

[36] Confidence here refers exclusively to a statement of perceived certainty made by the witness at the time of the initial identification. Confidence judgments are well known to be modified when an observer is later given access to other sources of information, such as other evidence and opinions about what transpired.

[37] Carolyn Semmler, John Dunn, Laura Mickes & John T. Wixted, *The Role of Estimator Variables in Eyewitness Identification*, 24 J. EXPERIMENTAL PSYCHOL.: APPLIED 400–15 (2018); Wixted & Wells, *supra* note 22, at 10–65; Wixted et al., *supra* note 22, at 515–26.

[38] Neil Brewer, Nathan Weber & Nicola Guerin, *Police Lineups of the Future?*, 75 AM. PSYCHOL. 76–91 (2020).

[39] This same perceptual scaling method is used in many practical applications, such as product marketing and preference testing, optometric assessment ("which is clearer: lens one or lens two?"), and hearing-aid evaluation.

[40] Many of these advances have been made possible by research grant support from the Laura and John Arnold Foundation, which issued a 2015 call for applications to fund research in areas highlighted by the NAS report.

[41] Roy S. Malpass & Patricia G. Devine, *Measuring the Fairness of Eyewitness Identification Lineups*, *in* EVALUATING WITNESS EVIDENCE 81–102 (S.M.A. LLOYD-BOSTOCK & B. R. CLIFFORD EDS. 1983); Gary L. Wells, Michael R. Leippe & Thomas M. Ostrom, *Guidelines for Empirically Assessing the Fairness of a Lineup*, LAW & HUM. BEHAV. 285–93 (1979).

[42] NATIONAL INSTITUTE OF JUSTICE TECHNICAL WORKING GROUP FOR EYEWITNESS EVIDENCE EYEWITNESS EVIDENCE: A TRAINER'S MANUAL FOR LAW ENFORCEMENT (2003).

[43] J. Efendi, M. Ihsan Zul & W. Yunanto, *Real Time Face Recognition Using Eigenface and Viola-Jones Face Detector*, 1 INT'L J. INFORMATICS VISUALIZATION 16–22 (2017); Tim Valentine, Michael B. Lewis & Peter J. Hill, *Face-Space: A Unifying Concept in Face Recognition Research*, 69 Q.J. EXPERIMENTAL PSYCHOL. 1996–2019 (2016); W. Zhao, R. Chellappa, P.J. Phillips & A. Rosenfeld, *Face Recognition: A Literature Survey*, 35 ACM COMPUT. SURV. 399–458 (2003); V. Blanz & T. Vetter (1999, July). *A Morphable Model for the Synthesis of 3D Faces*, in PROCEEDINGS OF THE 26TH ANNUAL CONFERENCE ON COMPUTER GRAPHICS & INTERACTIVE TECHNIQUES 187–94 (1999).

[44] C. Tredoux, *A Direct Measure of Facial Similarity and its Relation to Human Similarity Perceptions*, J. OF EXPERIMENTAL PSYCHOL.: APPLIED 180-93 (2002); M. Turk & A. Pentland, *Eigenfaces for Recognition*, 3 J. OF COGNITIVE NEUROSCIENCE 71–86 (1991).

[45] *See, e.g*., N. Brewer & G. Wells, The Confidence-Accuracy Relationship in Eyewitness Identification: Effects of Lineup Instructions, Foil Similarity, and Target-Absent Base Rates. J. OF EXPERIMENTAL PSYCHOL.: APPLIED 11-30 (2006).

[46] Gepshtein & Albright (2019), *supra* note 30; Sergei Gepshtein, Yurong Wang, Fangcho He, Dinh Diep, Thomas D. Albright, *A Perceptual Scaling Approach to Eyewitness Identification* (under review); Tim Valentine, Michael B. Lewis & Peter J. Hill, *Face-Space: A Unifying Concept in Face Recognition Research,* 69 Q.J. EXPERIMENTAL PSYCHOL. 1996–2019 (2016); Nancy Hirschberg, Lawrence E. Jones & Michael Haggerty, *What's in a Face: Individual Differences in Perception*, 12 J. RES. & PERSONALITY 488–99 (1978).

[47] Toni P. Saarela & Michael S. Landy, *Integration Trumps Selection in Object Recognition*, 25 CURRENT BIOLOGY 920 (2015).

[48] Jiayu Zhan, Oliver G. Garrod, Nicola van Rijsbergen & Philippe Schyns, *Modelling Face Memory Reveals Task-Generalizable Representations*, 3 NATURE HUM. BEHAV. 1–10 (2019).

[49] Ryan J. Fitzgerald, Heather L. Price & Tim Valentine, *Eyewitness Identification: Live, Photo, and Video Lineups*, 24 PSYCHOL. PUB. POL'Y LAW 307–25 (2018).

[50] IACP LAW ENFORCEMENT POL'Y CTR., EYEWITNESS IDENTIFICATION: MODEL POLICY CONCEPTS & ISSUES (2016), https://www.theiacp.org/resources/policy-center-resource/eyewitness-identification.

[51] U.S. DEP'T JUST., EYEWITNESS IDENTIFICATION PROCEDURES FOR CONDUCTING PHOTO ARRAYS (2017), https://www.justice.gov/file/923201/download.