The criminal justice system was designed to find and punish actual perpetrators guilty of the crimes of which they are accused. Questions and claims of innocence were rarely examined and, until recently, were generally dismissed out of hand. However, the advent of DNA-based exonerations some 30 years ago brought to light an alarming number of wrongful convictions that have called the system’s mission and methods into question.

Nevertheless, it is difficult to establish how often innocent defendants are found guilty, a point made repeatedly in the literature (e.g., Norris et al., 2017). Based on Norris et al. (2017) and other registry projections (e.g., National Registry of Exonerations) investigating samples of DNA-based exonerations, we estimate the frequency at about 5% in the United States. But surely this is a conservative estimate, as it is extrapolated from known exonerations and excludes wrongful convictions that have not been overturned. In any event, the National Registry of Exonerations reports that nearly 2,800 wrongful convictions have been reversed since 1989 (Norris et al., 2017). Fortunately, these shockingly high numbers have served as a clarion call to actively address the serious consequences of wrongful incarceration—which include not only the punishment of innocent people but citizens’ mistrust of the system and the failure to jail real perpetrators, who may go on to commit additional violent crimes.
As of 2018, work by the Innocence Project had led to 351 exonerations. Mistaken identification contributed to erroneous convictions in 239 (68%) of the cases. (The percentages sum to more than 100% because multiple causes contributed to many exonerees’ wrongful convictions.)

In 2016, we initiated a program of research exploring factors that contribute to wrongful convictions. Given our long-standing interest in memory and cognition, especially in the arena of eyewitness recollection and testimony, we were focused on memory factors that led juries and judges to reach guilty verdicts for defendants who later were determined to be not guilty. The basis for our research was a data set of such cases taken on by the Innocence Project, an initiative that strives to exonerate the wrongly convicted through DNA testing and criminal justice reforms. A large body of findings in the lab and the field has shown that eyewitnesses’ memory for perpetrators is subject to the same frailties and biases that typify recall of events and daily activities (Lindsay et al., 2007; Bialer et al. 2021). Although in some circumstances, mistaken identifications may be just as likely as correct identifications, lineups and showups (one-person “lineups”) are regularly used by investigators because there is no physical trace evidence in a large majority of cases (Wells, 1995; Wells & Loftus, 2003). In such cases, memory is the evidence!

Eyewitness lineup procedures at their core produce memory evidence which, as with any recollection, is subject to error and contamination. Though our research has a U.S. focus, the problem of wrongful convictions is global in scope. As an aside, we have collected international data (Toglia et al., 2018) in which we compared exonerations in the United States (N = 351), other Western countries (N = 900), and non-Western countries (N = 595) to further pinpoint factors responsible for convicting the innocent. Key results indicated that the leading cause of erroneous convictions outside of the United States was government misconduct (29%), followed by eyewitness memory issues (25%). To learn more about eyewitness identification guidelines worldwide, we recommend consulting Fitzgerald et al. (2021).
Psychological science in reforms

The impetus for reforming eyewitness-identification procedures was a seminal 1978 article by social psychologist (and APS James McKeen Cattell Fellow) Gary Wells distinguishing between estimator variables and system variables in criminal investigations. Following this classification, hundreds of laboratory studies have demonstrated the negative impact these variables may have, individually and collectively, on eyewitness memory and identification accuracy. Estimator variables are those present during a crime that could affect a witness’s accuracy (e.g., the witness’s stress level; the race of the perpetrator and the witness; and conditions that could affect the witness’s view, such as distance and lighting). They reduce the witness’s ability to encode and store the event (Wells & Olson, 2003) but are outside of the control of the legal system. Conversely, system variables (e.g., lineup instructions; selection of “fillers,” or people presented along with the suspect in lineups; and use/non-use of a blind administrator) can be controlled and modified by investigators. Using this partitioning scheme, Wells suggested new strategies for researching and reforming eyewitness-identification procedures in ways that could impede wrongful identification and, in turn, reduce the conviction of innocent individuals. His paper was an immediate game changer: In the 1980s, researchers armed with Wells’s distinction published a number of important laboratory and field studies showing that several lineup reforms (e.g., informing witnesses that suspects may or may not be in the lineup) reduced the identification of innocent suspects. Wells’s initial classification continues to have a tremendous influence on both methods and theory in the science of eyewitness lineups. More recently, it has also informed important strides in policy and reform (see Smith et al., 2021, for an in-depth review).

A large body of findings in the lab and the field has shown that eyewitnesses’ memory for perpetrators is subject to the same frailties and biases that typify recall of events and daily activities.

By the start of the 1990s, reform-oriented research by social and cognitive psychologists was growing exponentially and drawing increased attention in the criminal justice system. The timing of this growth was fortunate for the legal community. As framed by Toglia, Lampinen, and Smith (2021), “the criminal justice system found itself in a state of crisis” (p. 6) with the introduction of polymerase-chain-reaction DNA testing in 1984, followed shortly by the discovery of a disturbing number of innocent persons jailed (and, in some cases, sent to death row to await execution) for serious crimes they did not commit. These findings were a major force behind the launch of the Innocence Project in 1992 by Barry Scheck and Peter Neufeld.

Paralleling the genesis of the Innocence Project was the public’s increased awareness of miscarriages of justice, as portrayed through both news stories and TV crime dramas such as NYPD Blue and Law & Order. More recently, films such as Just Mercy and Conviction, both from 2019, and documentary series on Netflix and other streaming services (e.g., The Innocence Files, The Innocent Man) have focused on wrongful convictions involving actual events and people. These real-life stories have highlighted systemic problems within a judicial system that traditionally has focused on guilt. Only within the past decade has innocence been thrust into the public eye and national discourse.

Why the innocent are convicted
The criminal justice system treats physical evidence as items to be preserved and protected because they could be contaminated. Unfortunately, and inconsistent with psychological scientists’ recommendations, the system’s approach to the collection and preservation of memory evidence is qualitatively different (Wells & Loftus, 2003). Compared to the protection of crime scene evidence, investigators receive less formal training in the security of testimonial evidence obtained by interviewing witnesses concerning their memory for the crime and the suspect. Even though police interrogators typically have considerable experience in interviewing suspects and witnesses, they have less experience with issues regarding the vulnerability of human memory to suggestive information (see Loftus et al., 1978, for a seminal demonstration) that taints the collection of remembrances.

Given this distinction and our desire to limit contamination due to multiple causes contributing to wrongful convictions, for the purposes of this article we focused on memory errors involved in misidentification cases. Our first qualitative step, archival analyses, required homing in on mistaken-identification cases within the Innocence Project, which attributes wrongful convictions to five additional causes beyond eyewitness misidentification: unvalidated or improper forensic science, false confessions or admissions, government misconduct, informants (jailhouse prisoners who are often compromised), and inadequate defense.

Following our development of an initial coding scheme, we analyzed cases for 60 variables across five different categories (Toglia et al., 2017):

- Variables known to increase eyewitness inaccuracies during encoding (estimator variables) and at retrieval (system variables)
- Legal safeguards (e.g., expert testimony, presence of an attorney at the lineup)
- Suspect characteristics (e.g., race, juvenile status, mental disability)
- Case characteristics (e.g., conviction state, type of crime)

Determining the presence or absence of these factors is key to understanding their individual and combined roles in exoneration cases as well as the utility of recommended reforms.

We quickly realized that crucial data on these factors were missing from the Innocence Project website. Needing more data for an adequate archival study, we gained access to the Innocence Record, a database that houses documents concerning each exoneree’s conviction, including motions and transcripts from trials, such as witnesses’ testimonies and judges’ instructions.

The Innocence Record revealed 254 cases wherein the primary cause of conviction was eyewitness misidentification. We then narrowed the set to 104 cases in which mistaken identification was the sole factor for conviction—what we refer to as “pure” misidentification cases. Finally, we removed any profiles with little to no associated archival documents, reducing the final data set to 57 “pure” cases. Transcripts and other accompanying documents related to those cases provided a rich database that allowed us to use our five classification categories and expand our coding scheme to 123 variables.
### System Variables in Initial ID Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Single: $n = 26$ (28.9%)</th>
<th>Multiple: $n = 51$ (85.67%)</th>
<th>None/missing: $n = 13$ (14.4%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First ID test*</td>
<td>Photo array: $n = 43$ (47.8%)</td>
<td>Mugbook: $n = 6$ (6.7%)</td>
<td>Live lineup: $n = 10$ (11.1%)</td>
</tr>
<tr>
<td></td>
<td>Showup: $n = 14$ (15.6%)</td>
<td>Sketch: $n = 6$ (6.7%)</td>
<td>In-court ID: $n = 3$ (3.3%)</td>
</tr>
<tr>
<td></td>
<td>No ID test/missing: $n = 18$ (20%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lineup instructions for first ID*</td>
<td>Biased: $n = 12$ (22.6%)</td>
<td>Unbiased: $n = 28$ (52.8%)</td>
<td></td>
</tr>
<tr>
<td>Lineup administration for first ID*</td>
<td>Single-blind: $n = 38$ (71.7%)</td>
<td>Double-blind: $n = 10$ (18.9%)</td>
<td></td>
</tr>
<tr>
<td>Lineup presentation for first ID*</td>
<td>Simultaneous: $n = 37$ (69.8%)</td>
<td>Sequential: $n = 16$ (30.2%)</td>
<td></td>
</tr>
<tr>
<td>Video recording of first lineup*</td>
<td>Recorded: $n = 5$ (9.4%)</td>
<td>Not recorded: $n = 24$ (45.3%)</td>
<td>Unsure: $n = 24$ (45.3%)</td>
</tr>
</tbody>
</table>

### Legal Safeguards

<table>
<thead>
<tr>
<th>Variable</th>
<th>Present: $n = 2$ (3.5%)</th>
<th>Not present: $n = 35$ (61.4%)</th>
<th>Unsure: $n = 20$ (35.1%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attorney presence at lineup</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motion to suppress filed</td>
<td>Yes: $n = 27$ (47.4%); 6 (22.2%) granted</td>
<td>No: $n = 22$ (38.6%)</td>
<td>Unsure: $n = 8$ (14%)</td>
</tr>
<tr>
<td>Eyewitness expert testimony</td>
<td>Present: $n = 9$ (15.8%)</td>
<td>Rigorous: $n = 20$ (35.1%)</td>
<td>Standard: $n = 33$ (57.9%)</td>
</tr>
<tr>
<td>Cross-examination of detective</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judges’ instructions</td>
<td>Standard: $n = 19$ (33.3%)</td>
<td>Eyewitness-specific: $n = 14$ (24.6%)</td>
<td>Unavailable: $n = 14$ (24.6%)</td>
</tr>
</tbody>
</table>

### Estimator Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Other-race: $n = 25$ (43.9%)</th>
<th>Same-race: $n = 27$ (47.4%)</th>
<th>Unsure: $n = 5$ (8.2%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other-race ID</td>
<td>Present: $n = 42$ (73.7%)</td>
<td>Not present: $n = 14$ (24.6%)</td>
<td>Unsure: $n = 1$ (1.8%)</td>
</tr>
<tr>
<td>Weapon</td>
<td>Present: $n = 42$ (73.7%)</td>
<td>Not present: $n = 14$ (24.6%)</td>
<td>Unsure: $n = 1$ (1.8%)</td>
</tr>
<tr>
<td>Disguise</td>
<td>Disguised: $n = 5$ (8.8%)</td>
<td>Not disguised: $n = 47$ (82.5%)</td>
<td>Unsure: $n = 5$ (8.8%)</td>
</tr>
<tr>
<td>Lighting</td>
<td>Well-lit: $n = 28$ (49.1%)</td>
<td>Dark: $n = 22$ (38.6%)</td>
<td>Unsure: $n = 7$ (12.3%)</td>
</tr>
</tbody>
</table>
The table above shows the frequency of factors related to system variables, legal safeguards, and estimator variables in an archival analysis of 57 eyewitness misidentification cases.

The table at left provides a snapshot of the findings from these 57 cases. It is notable that the majority involved innocent suspects identified in multiple selection procedures. The use of multiple identification tests varies widely by jurisdiction, but more troubling is the increased risk of false identification resulting from repeated procedures (Steblay & Dysart 2016). That pattern in initial identifications, coupled with biased pre-lineup instructions in which the witness felt compelled to choose, is concerning, especially given that attorneys were rarely present during lineup administrations.

Another item that drew our attention was the more frequent use of simultaneously administrated lineups, in which lineup members or their photos are presented collectively, compared to sequential lineups, in which each person or photo is shown individually. The latter, it’s worth noting, has resulted in fewer misidentifications in lab studies (Cutler & Penrod, 1988), a finding supported by meta-analysis (Steblay et al., 2001).

Though the debate over the superiorit of sequential lineups continues (Kaesler et al., 2020; Steblay et al., 2011), theories of memory provide some perspective. One such position is fuzzy-trace theory (FTT). Advanced by Reyna and Brainerd (1995), FTT proposes that verbal and visual information is encoded in two types of independently developed memory traces. One results from processing verbatim aspects and contains exact, detailed information (e.g., a person’s specific facial and physical characteristics). The other represents the gist of the information, or general characteristics (e.g., a person’s gender, hair color, and approximate height). Both verbatim and gist representations contribute to accurate memory, but in explaining inaccuracies such as misidentifications, the focus is on gist processing. Returning to the comparison of lineup procedures, FTT predicts that simultaneous lineups provide considerable competing gist cues, obscuring differences among lineup options. Sequential lineups, by contrast, may result in more accurate identifications because they alternate a good verbatim cue (in the form of the actual perpetrator) with good gist cues, making the recollective differences more apparent.

Reliance on gist is also relevant to understanding the “own-race bias effect,” in which witnesses identify suspects of their own race better than suspects of other races. FTT argues that decreased familiarity with other races may increase interference from competing gist cues (Meissner & Brigham, 2001). In the entire Innocence Project database, 42% (375 post conviction DNA exonerations) of all cases involved cross-race misidentifications, and 44% (25) of the pure cases were cross-race mistaken identifications. Notably, each of those 25 cases involved a White witness. In 23 cases, the White witness misidentified a Black suspect; the other two misidentified suspects were Hispanic.

Referring again to the table, the estimator variables in the bottom panel include data regarding the frequency of crimes committed with a weapon (73.7%). This raises the specter of the “weapon-focus effect,” in which witnesses to a crime involving a weapon are much more likely to misidentify a suspect (Pickel, 1999). In FTT terminology, focused attention produces strong verbatim memories of the weapon but only gist impressions of the assailant’s face.

Moving from postdiction to prediction
Though other theories could be discussed, our current purpose is not to compare explanatory positions. Rather, FTT is our preferred approach to introduce postdiction, or a “looking back” strategy, which in the present context translates to leveraging theoretical expectations about what patterns should be evident in archival searches of wrongful-conviction cases.

This novel strategy is significant in a number of ways. In particular, it can improve the ecological validity of eyewitness research. Dissenters often question the admissibility of testimony based on that research and the generalizability of the findings because the methods used to test eyewitness factors do not parallel actual eyewitness situations (Kone?ni & Ebbesen, 1986). The courts have traditionally agreed with this argument, excluding expert testimony on the reliability of eyewitness memory on the grounds that it is not based on sufficiently established science (United States v. Amaral, 1973) or does not provide information beyond the jurors’ common sense (Schmechel et al., 2006). Arguments about the established science have largely abated, thanks to the Daubert standards, a set of guidelines on scientific testimony emanating from a Supreme Court case (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993), and more recent court decisions that have continued to offer corrective strategies, such as shifting the burden of proof to prosecutors to show that an eyewitness’s identification is sufficiently reliable (State v. Lawson, 2012) and the use of case-specific jury instructions to help jurors evaluate identification evidence (State v. Henderson, 2011). Nevertheless, external validity remains a concern. Using archival data is critical for reducing the discrepancy between laboratory simulations and real-world cases. By identifying system and estimator variables, legal safeguards, and case characteristics in documented exoneration cases, we can design new experimental paradigms and studies to argue and examine the factors most prevalent in actual cases of erroneous conviction due to mistaken identification. Materials carefully selected from real-world cases could also be used to design laboratory experiments modeled on actual crimes and trials.

**An experimental approach to informing reform**

Archival analyses from the Innocence Project and Innocence Record databases provide a glimpse into the benefits of accessing trial transcripts and witness testimony to understand factors (estimator and system variables) contributing to erroneous convictions, as well as to achieve further reform via existing or new legal safeguards. To date, analyses of DNA exoneration cases have isolated the impact of one factor, confidence (for a review, see Berkowitz et al., 2020), and one category, estimator variables (for a review, see Giacona et al. 2021), and have only been used for descriptive purposes. We endorse a new classification of study, which we call innocent-conviction research, to inform experimental designs. For example, the impact of eyewitness cross-examination as a legal safeguard against misidentification is well established in the literature (Berman & Cutler 1996). But laboratories have yet to explore how jurors assess the validity and accuracy of an identification when rigorous cross-examination strategies highlight differences between the identification procedures administered by a lead detective using recommended best-practice guidelines. This factor was present across all 57 pure misidentification cases in the Innocence Record database. We expect that as more studies emanate from archival descriptions, their findings will inform the further mining of archival databases, resulting in a symbiotic relationship between archival data and empirical research that allows for a more comprehensive understanding of wrongful convictions. Such an understanding should be extremely beneficial in further advancing criminal justice reform, expanding best practices (Wells et al., 2020), and ultimately restoring faith in the criminal justice system.
Perhaps the most important potential impact of this looking-back strategy is the furthering of communication between researchers and those working on behalf of erroneously convicted persons. Already, a significant development within the criminal justice system is the establishment of conviction review units (CRUs), which evaluate convictions in some jurisdictions to identify potential prosecutorial errors as an option to the traditional path of requesting an appeal post-conviction. Typically housed in district or state attorneys’ offices, CRUs carry out an extrajudicial examination of the facts in erroneous cases in which convicted defendants’ claims of innocence are highly plausible. CRUs also work on reforms designed to prevent unwarranted convictions.

It is not our purpose to detail here how CRUs function, their processes, or the many challenges they face (see Hollway, 2016, for an in-depth review and set of recommendations). Rather, our interest is to recognize important steps the criminal justice system has taken toward self-correction via CRUs—a marked contrast to prosecutors’ tendency to reject the possibility of either a flawed prosecution or actual innocence. CRUs’ entry into the role of aiding innocent people to be released from prison in 2007 has significantly increased the number of successful reversals of injustice. We call for researchers to work with CRUs, and vice versa, in using real-world archival data and identifying additional factors that need empirical study. Such efforts would produce a two-way street—a connection between the lab and the field—and would complement the model of a research agenda guided by postdiction, employed to test experimental predictions stemming from archival findings (e.g., our symbiosis argument), and having come full circle with fresh examinations of archival data repositories.

Finally, we urge law enforcement to improve record keeping so as to reduce the number of missing documents that CRUs and innocence commissions seek. Archival searches have allowed us to strengthen both the methodological strategies and rigor of research—but our review of the Innocence Project and Innocence Record files highlighted the need for a uniform organization of case files to improve accessibility and facilitate data mining. This will result in more incisive theoretical postdictions, greater clarity in proposing policy reforms and recommendations, and increased testing of experimental predictions. The details of wrongful convictions, regardless of their causes, also present opportunities to design stronger, ecologically valid “trial transcript” experiments, wherein participants read a detailed narrative (transcript) of a criminal court case, by modeling crime scenarios on actual Innocence Project and Innocence Record cases.

For generations, wrongful verdicts have had serious consequences that require addressing. Psychological science can help repair the public’s trust in the criminal justice system and eliminate threats to equal justice.

Feedback on this article? Email apsobserver@psychologicalscience.org or scroll down to comment.

References


Bialer, D. M., Reyna, V. F., & Brainerd, C. J. (2021). False memory: What are the effects, how does Fuzzy-Trace Theory predict them, and how does this matter for eyewitness testimony? In A. M. Smith,


United States v. Amaral, 488 F 2d 146 (9th Cir. 1973).


