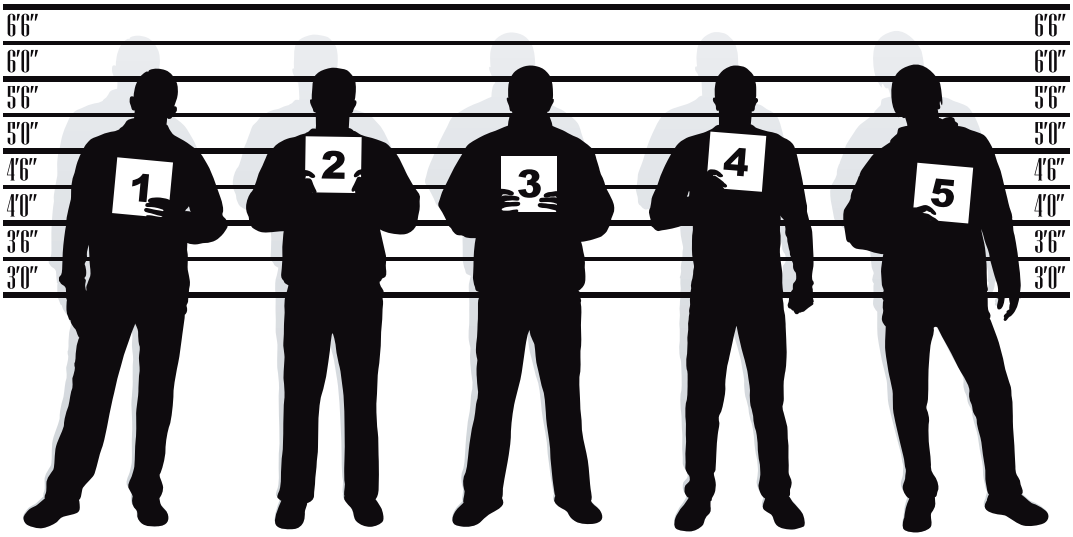
A Test of the Simultaneous vs. Sequential Lineup Methods

An Initial Report of the AJS National Eyewitness Identification Field Studies

GARY L. WELLS, NANCY K. STEBLAY, and JENNIFER E. DYSART



AMERICAN JUDICATURE SOCIETY SM

A Test of the Simultaneous vs. Sequential Lineup Methods

An Initial Report of the
AJS National Eyewitness Identification Field Studies

GARY L. WELLS, PhD
American Judicature Society
Center for Forensic Science & Public Policy
Iowa State University

NANCY K. STEBLAY, PhD
Augsburg College

JENNIFER E. DYSART, PhD
John Jay College of Criminal Justice

This study was supported by grants from the Open Society Foundations, the Laura and John Arnold Foundation, and the JEHT Foundation.
In addition, Nancy K. Steblay's work as principal investigator for the Tucson site was supported in part by a grant from the National Institute of Justice.

Seth S. Andersen
Executive Director

American Judicature Society
The Opperman Center
2700 University Avenue
Des Moines, Iowa 50311
www.ajs.org

Founded in 1913, the ***American Judicature Society*** is an independent, non-partisan, membership organization working nationally to protect the integrity of the American justice system. AJS's diverse and broadly based membership – including judges, lawyers, and members of the public – promotes fair and impartial courts through research, publications, education, and advocacy for judicial reform. The work of AJS focuses primarily on judicial ethics, judicial selection, access to justice, criminal justice reform, and the jury system.

Eyewitness Identification Field Studies Partnership

Lead Scientists

GARY L. WELLS, PhD.
Distinguished Professor of Psychology, Liberal Arts & Sciences
Wendy & Mark Stavish Chair in Social Sciences
Iowa State University

Director of Social Sciences
American Judicature Society Center for Forensic Science & Public Policy

NANCY K. STEBLAY, PhD.
Professor of Psychology
Augsburg College

JENNIFER E. DYSART, PhD.
Associate Professor of Psychology
John Jay College of Criminal Justice

American Judicature Society

Seth S. Andersen
Executive Director

Danielle Mitchell
Project Manager
American Judicature Society
Center for Forensic Science & Public Policy

Christine Mumma
Executive Director
North Carolina Center on Actual Innocence
Board of Directors
American Judicature Society

Innocence Project[1]

Barry Scheck
Co-Director

Maddy deLone
Executive Director

Karen Newirth
Eyewitness Identification Litigation Fellow

Ezekiel Edwards
Staff Attorney on Eyewitness Identification

Emily West
Research Director

Audrey Levitin
Director of Development

Police Foundation

Hubert Williams
President

Dr. Karen Amendola
Chief Operating Officer
Research, Evaluation & Professional Services

Meghan Slipka
Research Associate
Research, Evaluation & Professional Services

Edwin Hamilton
Professional Services Director

Center for Problem Oriented Policing

Michael Scott
Director

1 The Innocence Project self-funded its participation in the Eyewitness Identification Field Studies, which included significant contributions of professional staff time, travel and OTP costs

Acknowledgements

The authors and Eyewitness Identification (EWID) partnership offer special thanks to the following individuals for their advice and support during the development and implementation of the AJS National Eyewitness Identification Field Studies.

Hon. Janet Reno
Fmr. Attorney General of the United States

James M. Doyle
Fmr. Consulting Director
Center for Modern Forensic Practice at
John Jay College of Criminal Justice
Attorney, Carney & Bassil

Michael Garner
Director
PS&J OneSolution Development
SunGard Public Sector, Inc.

The authors and EWID partnership offer special thanks to the following police departments, law enforcement executives and personnel, district attorneys, and legal professionals for their participation and their commitment to ensuring that the EWID Field Studies were a success. Without their partnership and support, the Field Studies would not have been possible.

Austin Police Department
Art Acevedo
Chief of Police

David Carter
Assistant Chief of Police, Chief of Staff

Patti Robinson
Assistant Chief of Police, Headquarters Bureau

Julie O'Brien, Commander
Headquarters Bureau – Violent Crimes

Mark Spangler, Lieutenant
Strategic Intelligence –ARIC

Troy Gay, Commander
Strategic Planning

Pat Cochran, Lieutenant
Strategic Intelligence

Arthur Arevalo, Supervisor
Police Technology Unit

JJ Schmidt, SPO
Police Technology Unit

Simone Graboskie, SPO
Police Technology Unit

Eric Burnsed, SPO
Police Technology Unit

Brenda Barnard
City of Austin
IT Project Manager

Shaun Davis
City of Austin
Senior Network Administrator

Travis County District Attorney's Office
Rosemary Lehmberg
District Attorney

John Neal
First Assistant District Attorney

Bryan Case
Director of Appellate Division

Buddy Meyer
Director of Trial Bureau

Charlotte-Mecklenburg Police Department
Rodney Monroe
Chief of Police

Darrel W. Stephens
Chief of Police (ret.)

Kevin Wittman, Major (ret.)
Criminal Investigations Bureau

Doug Gallant, Major
Southwest Service Area

Paul Zinkann III, Major
Southwest Service Area/Criminal Investigations Bureau

Kevin Barbee, Sergeant
Criminal Investigations Bureau/Armed Robbery Unit

Kenneth Clark, Sergeant (ret.)
Supervisor Armed Robbery Unit

Thanks also to Brian Cutler, *Professor of Psychology, University of Ontario Institute of Technology,* and Lori Van Wallendael, *Associate Professor, University of North Carolina at Charlotte,* for their assistance at the Charlotte-Mecklenburg Police Department site.

Mecklenburg County District Attorney's Office

R. Andrew Murray
District Attorney

Peter Gilchrest III
District Attorney (ret.)

Jay Ashendorf
Assistant District Attorney

Bart Menser
Assistant District Attorney

Marsha Goodenow
Fmr. Assistant District Attorney

Sean Poll
Assistant District Attorney

Tucson Police Department

Roberto Villaseñor
Chief of Police

Richard Miranda
Chief of Police (ret.)

Kermit Miller
Chief of Police (ret.)

Sharon Allen
Deputy Chief

John Stamatopoulos, Captain
Specialized Response Division

John Neeley
ID Superintendent

Thanks also to Nicholas Fernandez, PhD Candidate, *University of Arizona,* for his assistance at the Tucson Police Department site.

Tucson City Prosecutor's Office

Lisa Judge
Principal Assistant City Attorney
TPD Legal Advisor

Pima County Attorney's Office

Barbara LaWall
Pima County Attorney

Kathleen Mayer
Deputy County Attorney's Office

Arizona Attorney General

John Evans
Fmr. Assistant Attorney General for Arizona

Pima County Public Defender's Office

Robert J. Hirsh
Pima County Public Defender

University of Arizona

Gabriel J. (Jack) Chin
Fmr. Chester H. Smith Professor of Law
University of Arizona College of Law

Roger Hartley
Fmr. Assistant Professor
School of Public Administration & Policy
University of Arizona Eller College of Business & Public Administration

Thank you to Michael Bloom, Esq., *Criminal Defense Attorney,* for his advice on the Tucson Police Department site.

San Diego Police Department

William Lansdowne
Chief of Police

James Collins, Captain
Investigations

Kevin Ammon, Lieutenant
Robbery Unit

Karen Goodman
Crime Analyst Supervisor

Daniel Newland
Information Systems Administrator

Kimberly Glenn
Information Systems Analyst Manager

Hong Thai
Information Systems Analyst III

San Diego County District Attorney's Office
Bonnie Dumanis
District Attorney

Tia Quick
Deputy District Attorney
SPD Legal Liaison

The authors and EWID Partnership offer special thanks to the following individuals for their support and input on this project:

William Berger
Chief of Police (ret.)
Palm Bay Police Department

Dr. Steven E. Clark
Professor of Psychology
University of California – Irvine

Hon. George "Woody" Clarke
San Diego Superior Court

Keith Findley
Co-Director, Wisconsin Innocence Project
Clinical Professor of Law, University of Wisconsin

Dr. Roderick C.L. Lindsay
Professor of Psychology
Queens University

Dr. Elizabeth F. Loftus
Distinguished Professor Psychology
University of California – Irvine

Dr. Otto H. MacLin
Associate Professor
University of Northern Iowa

Dr. Roy S. Malpass
Professor Emeritus of Psychology
University of Texas at El Paso

Dr. Steven D. Penrod
Distinguished Professor Psychology
John Jay College of Criminal Justice

Thomas Sullivan
Attorney
Jenner & Block

Jim Tooney
Attorney
Charlotte, NC

Tom Tyler
Professor of Psychology
New York University

A Test of the Simultaneous vs. Sequential Lineup Methods

An Initial Report of the AJS National Eyewitness Identification Field Studies

Executive Summary

The significant role that mistaken eyewitness identifications have played in convictions of the innocent has led to a strong interest in finding ways to reduce eyewitness identification errors. Psychological scientists have been conducting laboratory studies on this problem for over 30 years and have proposed a number of possible reforms to the procedures used in conducting lineups. Most of the proposed reforms, including the critical requirement of double-blind administration (the administrator does not know the identity of the suspect), have not been considered controversial in principle and many jurisdictions across the United States have adopted them. The use of a double-blind (DB) sequential rather than a DB simultaneous lineup procedure, however, has engendered controversy, a controversy that has unnecessarily held back the adoption of non-controversial reforms in many jurisdictions.

The sequential lineup shows lineup members to the witness one at a time and asks the witness to make a decision on each one before showing the next one, whereas the traditional simultaneous lineup shows the witness all lineup members at once. Controlled laboratory experiments consistently show that the DB sequential procedure results in a substantial reduction of mistaken identifications and a much smaller reduction in accurate identifications. Overall, the DB sequential lineup produces a better ratio of accurate identifications to mistaken identifications than the DB simultaneous procedure. Nevertheless, in May of 2006, a highly publicized field study in Illinois, directed by the Chicago Police Department not only called into question the sequential/simultaneous laboratory findings but raised concerns as to whether eyewitnesses in controlled experiments were a good approximation for actual eyewitnesses to serious crimes, a large share of which are victim-witnesses. Specifically, the Illinois study showed that the status quo method produced higher suspect identification rates and lower filler picks than did DB sequential lineups in two of the three cities that were tested. Lineup fillers are not suspects but instead are in the lineup to "fill it out" and create a fair procedure for the suspect. In a field experiment, the identification of fillers is the only witness response that can be definitively classified as an error.

The Illinois study was quickly rejected by scientists for several reasons. Principal among the reasons were (a) that this field study confounded the simultaneous/sequential variable with non-blind versus double-blind testing, (b) there was no random assignment of cases to lineup procedure and later evidence from the Evanston site indicated that the "tougher" cases (e.g., cross-race, longer delay from crime to lineup) were more likely to be assigned to the sequential than to the simultaneous procedure, and (c) some unknown number of filler identifications were not recorded for the simultaneous lineups. Consequently, in September of 2006, the American Judicature Society convened a gathering of eyewitness scientists, lawyers, prosecutors, and police in Greensboro, NC, who developed what has become known as the "Greensboro Protocol." The Greensboro Protocol was a set of guidelines for how to conduct a field experiment to test the simultaneous versus sequential issue and gather as much reliable data as possible

on witness and event variables (e.g., type of crime, presence of a weapon, cross- race event, viewing conditions, previous acquaintance with the culprit, sobriety of the witness), and the actual administration of the lineup itself (e.g., time between crime and lineup, quality of lineup, the witness's responses and statement of certainty). There was general agreement that the field study should feature a direct comparison of DB sequential and DB simultaneous procedures, true random assignment (the "gold standard" in scientific experiments), and the use of laptop computers.

The use of laptop computers for administering the lineup and recording the witnesses responses was believed to be an especially important tool for conducting eyewitness field experiments because it could: 1) Ensure procedures were administered according to protocol (e.g., voice and printed pre-lineup instructions presented in every instance in a uniform fashion); 2) Reliably record all responses of the witness (e.g., no selectivity in deciding whether to make a record of a filler identification or lack of an identification); 3) Permit all the photos in a lineup to be preserved as part of the electronic record and reviewed subsequently by judges, juries, and scientists; 4) Randomly assign witnesses to conditions (e.g., whether a sequential or simultaneous procedure would be used); 5) Randomly determine order of the photos within each lineup; 6) Precisely record how long it took a witness to make an identification; 7) Require police officers to record systematically witness and event variables before the identification procedure was conducted; 8) Facilitate secure and contemporaneous recording of eyewitness data into the electronic information platforms of police departments; and 9) Enhance the confidence of prosecutors, judges, juries, and defense counsel that the eyewitness procedures were conducted fairly and in accordance with best practices. In short, there was an expectation that the design of this field study and the use of the laptop computers could produce a data set of unprecedented depth and detail beyond the sequential/ simultaneous question.

The field experiment was developed, sites were recruited, and funding from foundations was secured with the help of many individuals and organizations. The funding foundations were the Laura and John Arnold Foundation, the Open Society Foundations, and the JEHT Foundation. The American Judicature Society oversaw the project with Danielle Mitchell as the project manager. Partner organizations included the Innocence Project, the Police Foundation, and the Center for Problem-Oriented Policing.

Three scientists (Dr. Gary Wells, Dr. Nancy Steblay, & Dr. Jennifer Dysart) were intimately involved in the design, implementation, detective training, and analysis of data. Mike Garner of SunGard Public Sector, Inc. programmed the software for the laptop computers that ran the lineups. Many other scientists, police officers, prosecutors, defense lawyers, and judges contributed to the development of the Greensboro protocol and actual language that appears on the different screens shown to witnesses. The District Attorney's offices at each of the sites gave their full cooperation to the project and the police departments at each of the sites were extremely cooperative, helpful, and, of course, essential in getting this project completed. A more detailed set of acknowledgements is contained in an Acknowledgements section.

The field experiment was conducted in the Charlotte-Mecklenburg (NC) Police Department, the Tucson (AZ) Police Department, the San Diego (CA) Police Department and the Austin (TX) Police Department. For various reasons, most of the data came

from Austin and the samples from the other three sites were not large enough to test for differences across sites. Thus the data were collapsed across the four sites in this report. Cases for which the eyewitness had prior knowledge of the suspect (e.g., went to school together or in some other way were previously acquainted) were removed from the primary set of lineups as were those instances in which the lineup was not conducted using a double-blind procedure. The resulting "protocol-consistent" set of cases totaled 497 lineups ranging in seriousness from simple assault to murder. Analyses of the data indicated that random assignment to condition was highly successful (e.g., key witness and event variables were equally distributed across simultaneous and sequential procedures and positioning of the suspect was equally distributed across simultaneous and sequential procedures).

Results

The overall analysis of identification data found that the simultaneous procedure yielded an identification of the suspect for 25.5% of the lineups and the sequential procedure yielded an identification of the suspect for 27.3% of the lineups. Statistical analyses showed that the simultaneous versus sequential difference in rates of identifying the suspect was not statistically significant. In other words, the small difference in suspect identification rates is within the margin of error (mere chance) and should not be interpreted as a meaningful difference. An analysis of filler identification rates, however, found that the simultaneous procedure yielded 18.1% identifications of fillers and the sequential procedure yielded 12.2% identifications of fillers. Statistical analyses showed that the rate of filler identifications was a statistically significant difference using a conventional probability level of $p < .05$. In other words, statistical analyses indicated that there is less than a 5% probability that the 5.9% lower rate of filler identifications using the sequential procedure was due to chance. The rates of non-identification were 56.4% for the simultaneous and 60.5% for the sequential, a difference that is due to the lower rate of filler identifications for the sequential procedure. A closer analysis showed that 80.8% of the non-identifications were clear rejections ("no" to all photos) for the simultaneous procedure and 19.2% were "not sure" responses. In contrast, 53.5% of the non-identifications were clear rejections for the sequential procedure and 46.5% were "not sure" responses.

The results are consistent with decades of laboratory research showing that the sequential procedure reduces mistaken identifications with little or no reduction in accurate identifications. Deeper analyses will be conducted on many other aspects of the data, including the certainty statements of the eyewitnesses, which were audio recorded for all identifications. This could tell us, for example, whether mistaken identifications of fillers using one procedure versus the other are associated with more certainty. Analyses will also be conducted on whether the witness was a victim-witness or a bystander witness, whether the identification was same-race versus other-race, and numerous other measured variables to see if superiority of the sequential procedure is restricted to certain types of circumstances. Scholarly articles will be published in refereed scientific journals that report these deeper analyses. Also, although filler identifications are clearly errors, identifications of the suspect might or might not be accurate identifications. Accordingly, a second phase of the research, led by the Police Foundation, is being conducted that follows up on these cases.

A Test of the Simultaneous vs.Sequential Lineup Methods

An Initial Report of the AJS National Eyewitness Identification Field Studies

Background

The identification of perpetrators from police lineups is an important tool in solving crimes and convicting guilty individuals. A lineup appears on the surface to be a simple, straightforward test. If the suspect is guilty, the eyewitness might be able to identify him; if he is not guilty the eyewitness will identify no one. Psychological scientists, who have been studying eyewitness identification evidence for over 35 years, however, have long questioned this view. Using simulated crimes, these researchers have noted that mistaken identification rates can be surprisingly high under some conditions. Particular interest has been taken by researchers in what are known as "system variables," which are variables that the justice system can control that increase or decrease the accuracy of eyewitness identification evidence[1]. By the late 1980s, eyewitness researchers had described several system-variable improvements or reforms that could increase the probative value of lineups, including the sequential lineup method[2].

Although the studies on eyewitness identification were well accepted in the scientific psychology community, it was not until the mid-1990s that the legal system began to take the issue of mistaken eyewitness identification more seriously. The impetus for this was a series of convictions of the innocent that were overturned using forensic DNA evidence. The Innocence Project in New York, launched by Barry Scheck and Peter Neufeld, marked the first systemic effort to test innocence claims of convicted prisoners using the new technique of forensic DNA testing. Forensic DNA testing represented the first scientific test in the history of the criminal justice system that could definitively prove actual innocence under certain conditions[3]. Not surprisingly, these turned out to be almost exclusively sexual assault and murder cases because those are the cases for which certain conditions tend to be met; DNA-rich biological evidence (semen and blood) was collected and preserved and can be clearly attributed to the perpetrator. For purposes related to the current report, it is noteworthy that 75% of the exonerations, which now number 273, were cases that involved mistaken eyewitness identifications[4]. However, most cases of wrongful conviction that are based on mistaken eyewitness identifications can never be discovered with DNA tests because the biological evidence was lost, destroyed, deteriorated, not collected, or not collected properly. Moreover, only a small fraction of eyewitness identification cases (estimated at less than 5%) have biological evidence that can be tested for purposes of possibly trumping or validating the eyewitness identification. Hence, forensic DNA testing can only test a small fraction of innocence claims in old cases and can screen out only a small fraction of mistaken identifications in current cases. The result is that the criminal justice system is still heavily dependent on eyewitness identification evidence and therefore improving the reliability of eyewitness identification evidence remains an important goal.

There are now hundreds of published studies that use simulated crimes followed by lineups[5] in which various conditions are systematically changed to see how those variations affect rates of mistaken and accurate identifications. This report refers to this body of literature as the lab studies. These lab studies are highly controlled in the sense that they vary only one factor at a time so that differences in eyewitness identification performance can be attributed to the factor that was varied rather than extraneous, uncontrolled factors. Importantly, because the researchers are the ones who created the witnessed event and the lineup, it is known with total certainty which member of

1. Wells, G. L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*, 1546-1557.
2. Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentations. *Journal of Applied Psychology, 70*, 556-564.
3. Scheck, B., Neufeld, P. & Dwyer, J. (2000). *Actual innocence.* New York: Random House.
4. As of July 29, 2011
5. The term lineup in this report is used to refer to either a photographic array or a live lineup.

the lineup, if any, is the actual perpetrator of the simulated crime. As a result, lab researchers can classify eyewitness identifications as having been accurate or mistaken without any doubt about the classification. Eyewitness scientists tend to prefer this lab-controlled experimental methodology for getting at cause-effect relations[6]. It is out of this lab-based research literature that most of the current literature of how to improve lineups was born[7].

Resistance to some reform ideas has understandably surfaced among some in the legal system. While some jurisdictions have embraced the reforms, others raise serious concerns about changing their current, long-standing practices based on lab studies that have not been fully tested in actual cases. There are several arguments for dismissal of the lab studies: Participants in an experiment are not eyewitnesses to actual crimes and are usually debriefed shortly after viewing a simulated crime that it was not an actual crime; the consequences for mistakes in research participants' lineup decisions are not as serious as in actual cases, perhaps leading to some mere guessing; participants in experiments are not experiencing the levels of stress and fear that many actual eyewitnesses experience, especially victim-witnesses; and, a large percentage of lab studies rely on college students as their subjects, which is unrepresentative of the typical witness in an actual case. Eyewitness scientists have countered these criticisms with a number of observations. For example, studies have found that whether participants believe the crime was real versus simulated does not matter to the results; controlled studies testing the role of stress and fear show that both serve to reduce accuracy; and lab studies comparing college students to other populations show that college students are the best witnesses of all groups. Accordingly, the eyewitness scientists argue, the lab studies might actually be *overestimating* rather than underestimating the accuracy of eyewitnesses in actual cases.

Despite arguments for and against the utility of the lab experiments in extrapolating to actual crimes and eyewitnesses, the value of testing some key ideas of eyewitness scientists using actual eyewitnesses to serious crimes is undeniable. This is especially true for the somewhat controversial idea of sequential lineups. The sequential lineup is one in which the witness views each lineup member one at a time and makes an identification decision on each before seeing the next lineup member rather than viewing all lineup members as a group (simultaneous). The sequential lineup was first tested in lab studies in 1985 and was predicted to be superior to the simultaneous method based on an emerging theory that eyewitnesses have a tendency to use relative judgments in making eyewitness identification decisions[8]. A relative judgment is one in which witnesses compare lineup members to one another and try to decide which one looks most like their memory of the perpetrator. Witnesses then have a propensity to select that person. The problem with relative judgment, according to the theory, is that someone will always look more like the perpetrator than the other members of the lineup, even when the lineup does not contain the perpetrator. A reliable effect, called the removal-without-replacement effect, was demonstrated in lab experiments in 1993 and has served as one of the core findings illustrating the relative-judgment process[9]. The effect simply shows that if the actual perpetrator is removed from a lineup and replaced with no one, a large share of eyewitnesses who would have picked the perpetrator tend to shift to another lineup member and identify that person rather than make no identification (even though they are clearly warned that the actual perpetrator might not be in the lineup).

The idea of the sequential lineup, therefore, was to prevent witnesses from merely comparing one lineup member to other lineups members (a relative judgment) and instead to compare each lineup member to their memory of the perpetrator and make an "absolute" judgment. Although relative judgments could still be made at one level with the sequential procedure (e.g., this person looks more like the perpetrator than the previous one), relative judgment should largely be blocked as long as the witness presumes that there still are (or might be) other lineup members to be viewed.

It is not the purpose of this report to review the lab studies of the sequential versus simultaneous procedure. Those who are interested can read the recent meta-analysis article that analyzed the extant literature and found that the sequential procedure produces a better ratio of accurate to mistaken identifications than does the simultaneous

6. Wells, G. L., & Penrod, S. D. (2011). Eyewitness identification research: Strengths and weaknesses of alternative methods. In B. Rosenfeld, & S. D. Penrod (Eds.), *Research methods in forensic psychology*. John Wiley and Sons, Hoboken, NJ.

7. Wells, G. L., Small, M., Penrod, S. J., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22,* 603-647.

8. Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14,* 89-103.

9. Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48,* 553-571.

procedure[10]. But, the meta-analysis also provides evidence that with the sequential procedure accurate identifications might be reduced by 8% even while mistaken identifications are reduced by 22%. It is this possibility of lower rates of accurate identifications that has made the sequential lineup controversial. Supporters of the sequential lineup tend to argue that the 8% loss in accurate identifications is the result of diminished rates of guessing or that the important figure is the ratio of accurate to mistaken identifications. Critics argue that a loss of accurate identifications is a serious cost and should give pause to any jurisdiction about switching to sequential lineups. Supporters counter that any mistaken identification not only puts the innocent at risk of wrongful conviction but also lets the guilty escape detection and, hence, the simultaneous lineup also lets the guilty go free when the witness mistakenly identifies someone else[11].

Furthermore, the "lab studies are just lab studies" argument keeps open the possibility that the sequential procedure does not work as well in actual cases when compared to the traditional simultaneous procedure as it does in the lab. By the same token, the sequential procedure might work even better in actual cases than it does in the lab when compared to the simultaneous procedure. Continued back and forth arguments will never resolve the question. Hence, there is a need to compare the two lineup techniques in actual criminal cases.

The difficulties with testing the sequential versus simultaneous question in actual cases are numerous but surmountable. The most important difference between the lab and actual cases is that the actual identity of the "perpetrator" is known with certainty in the lab but not – in most circumstances - in actual cases. When an eyewitness identifies a suspect in an actual case, we cannot presume that the suspect is in fact the guilty person. Accordingly, we will use the term "suspect" in this report to refer to the focus person in the lineup, a term that is meant to embrace other designations used by law enforcement such as "possible suspect" or "person of interest." In any case, the term suspect should not be confused to mean perpetrator or culprit. A procedure that produces more or fewer identifications of the suspect might or might not be the best procedure, depending on the proportion or mix of innocent versus guilty suspect identifications that each procedure produces.

Nonetheless, this limitation about the guilty status of a suspect in a lineup does not prevent us from concluding that filler identifications are clearly mistaken identifications. A filler is a known-innocent member of the lineup whose presence in the lineup is merely to "fill it out" and help safeguard an innocent suspect by spreading identification mistakes across people who will not be charged if they are identified. A clear requirement that the current research placed on the police department sites is that every lineup include only one suspect embedded among five fillers. This permitted a test of a central tenet of the sequential procedure, namely that it reduces filler identifications. In other words, filler identifications serve as a clear proxy index for the relative ability of the two procedures to reduce mistaken identifications.

Furthermore, in order to effectively test the two lineup procedures, it is necessary to use a strategy that can equalize the proportions of guilty (and innocent) suspects between the two compared procedures; whatever the true proportion of guilty suspects in these lineups, true random assignment to sequential and simultaneous lineups will distribute this factor evenly between the tested groups.

The Illinois Study

In 2006 a study was conducted in the Evanston, IL Police Department, the Joliet, IL Police Department, and two stations of the Chicago Police Department. The study's stated purpose was to compare a new procedure for conducting lineups, in particular the sequential double-blind method, to the traditional simultaneous non-blind procedure[12]. The research design, however, was problematic from the outset[13]. Eyewitness scientists had long argued that non-blind administration of lineups (i.e., the lineup administrator knows which lineup member is the suspect and which are fillers) will tend to lower filler identifications and raise identifications of the suspect via unintentional cues from the lineup administrator and, therefore, all lineups should use double-blind methods to ensure that the witness is

10. Steblay, N., Dysart, J. & 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,* 99-139.

11. As of July 29, 2011, the actual perpetrators have been discovered in 36% of the DNA exonerations involving mistaken identification. Thirty-one of the actual perpetrators were convicted of 77 violent crimes that they committed after the wrongful convictions involving mistaken identifications. These included 52 rapes, 17 murders, and 8 other violent crimes. [Source: WWW.innocence project.org]

12. Although there were two social science consultants involved in the data analysis, it is unclear that they played much role in the design. The design, procedure, interpretations, and conclusions were headed by an attorney with the Chicago Police Department.

making the identification based purely on his or her own memory. Yet the Illinois study always allowed the case detectives to administer their own lineups (non-blind) for the simultaneous procedure whereas the sequential procedure was always conducted double-blind. The results reported in the Illinois study indicated that the traditional non-blind lineups produced fewer filler identifications and more suspect identifications than did the double-blind lineups. But, of course, the non-blind lineups were simultaneous and the double-blind lineups were sequential, so it is unclear what caused this difference in witness identification decisions. Indeed, the social science concerns about non-blind lineup administration is that the case detective can unintentionally, and without awareness, influence the witness away from fillers and toward the suspect. Hence, it would be just as valid to interpret the results of the Illinois study as evidence that the non-blind lineup administrators influenced the witnesses' identifications as to interpret the difference as being due to the simultaneous versus sequential component of the study. Also, a later addendum report from the Illinois study acknowledged that some unknown number of filler identifications for the simultaneous procedure were not recorded in the results because the (non-blind) detective decided that the witness was not sure enough in the identification. Furthermore, none of the sites used random assignment to conditions, an essential requirement for a valid experiment. When the Evanston Police Department released the data from its portion of the study, analyses showed that the "tougher" identification cases (e.g., cross race cases, longer delay from the witnessed event to the time of the lineup[14]) were somehow assigned to the sequential lineups more often than to the simultaneous lineups[15].

Greensboro Meeting

Later in 2006, an eyewitness field study meeting was held in Greensboro, NC. Many of the top eyewitness identification scientists in the country along with lawyers, prosecutors, and law enforcement with expertise in eyewitness issues discussed what questions a field study on eyewitness identification could answer and what kinds of scientific controls were necessary to conduct a field experiment that would answer the questions. From these meetings came what has been called the "Greensboro Protocol." In effect, the Greensboro Protocol articulated the view that any field study would have to use double-blind lineup procedures in all conditions in order for the results to be accepted by the scientific community. Furthermore, the use of laptop computers for administering the lineup and recording the witnesses responses was needed to ensure that the procedures were administered according to protocol (e.g., voice and printed pre-lineup instructions presented in every instance in a uniform fashion) and that all responses of the witness were recorded (e.g., no selectivity in deciding whether to make a record of a filler identification or lack of an suspect identification). Furthermore, the use of the computer would permit all the photos to be preserved as part of the electronic record, the computer could randomly determine the order of the photos, the computer could randomly assign witnesses to conditions, and so on. Finally, there was a general view that the simultaneous versus sequential issue would be the important question to test, especially since the sequential procedure was already in use in several jurisdictions (e.g., New Jersey, Boston, and Minneapolis) and the Illinois study had created concern and confusion on the issue[16].

A Partnership to Conduct the Current Study

After the Greensboro meetings, a partnership developed between the American Judicature Society (represented at the time by Christine Mumma), the Innocence Project (Barry Scheck and Ezekiel Edwards), the Police Foundation in Washington, DC (Karen Amendola and Megan Slipka), the Center for Problem-Oriented Policing, and social scientists. A records management company in North Carolina (SunGard

13. The problem that was most apparent was the confounding of the simultaneous versus sequential variable with the non-blind versus double-blind variable. Not only eyewitness scientists, but also top social scientists who are not involved in the eyewitness research area concluded that the confound "has devastating consequences for assessing the real-world implications of this particular study" [see Schacter, D., Dawes, R., Jacoby, L. L., Kahneman, D., Lempert, R., Roediger, H. L., Rosenthal, R. (2007). Studying eyewitness investigations in the field. *Law and Human Behavior, 32*, 3-5.]

14. "Tougher" in this context means that witnesses tend to perform more poorly under these conditions according to published controlled studies.

15. Steblay, N. K. (2011). What we now know: The Evanston Illinois field lineups. *Law and Human Behavior, 35*, 1-12.

16. Extensive consideration was given by the partners to the question of whether to test the double-blind versus non-blind question as well. However, this idea was rejected for several reasons. The most important reason is that any better "performance" of witnesses in the non-blind compared to the double-blind conditions would simply raise the interpretational argument that the detectives were influencing witnesses away from fillers and toward suspects in the non-blind conditions. Such a finding would make any law enforcement agency participating in the study look bad. Furthermore, showing influence from detectives in the non-blind conditions would present problems for prosecutors in those cases. Alternatively, if no differences emerged between the double-blind and non-blind conditions, the lack of a difference might be attributed merely to the "observation" effect in which the behaviors of the investigators were unnaturally controlled due to their being observed in a study.

Public Sector, Inc.) with extensive experience with law enforcement agencies was added as a partner to develop and administer the software that would run the lineups on laptop computers and collect witness response data according to the protocol and needs of the study. SunGard Public Sector, Inc., was working with the Winston-Salem and Burlington (NC) police departments and had developed computer-based photo lineup software for these departments. Mike Garner, of SunGard Public Sector, Inc., had presented a version of this software at the Greensboro Conference. SunGard Public Sector, Inc. had the capability to create an interface between the laptop computers and the mainframe server in Charlotte-Mecklenburg (our first site) so that the results from each lineup could be uploaded after each lineup was completed. To ensure consistency with all sites, SunGard Public Sector, Inc. had to create a similar interface with each department that participated in the current study[17]. Four police departments, the Charlotte-Mecklenburg (NC) Police Department, the Tucson (AZ) Police Department, the San Diego (CA) Police Department, and the Austin (TX) Police Department, along with their respective prosecutor offices agreed to be sites for conducting the study. Funding was provided by the Open Society Foundations, the Laura and John Arnold Foundation and the JEHT Foundation. Throughout the project, many people made important contributions (see Acknowledgements).

Methods Used to Collect the Field Study Data

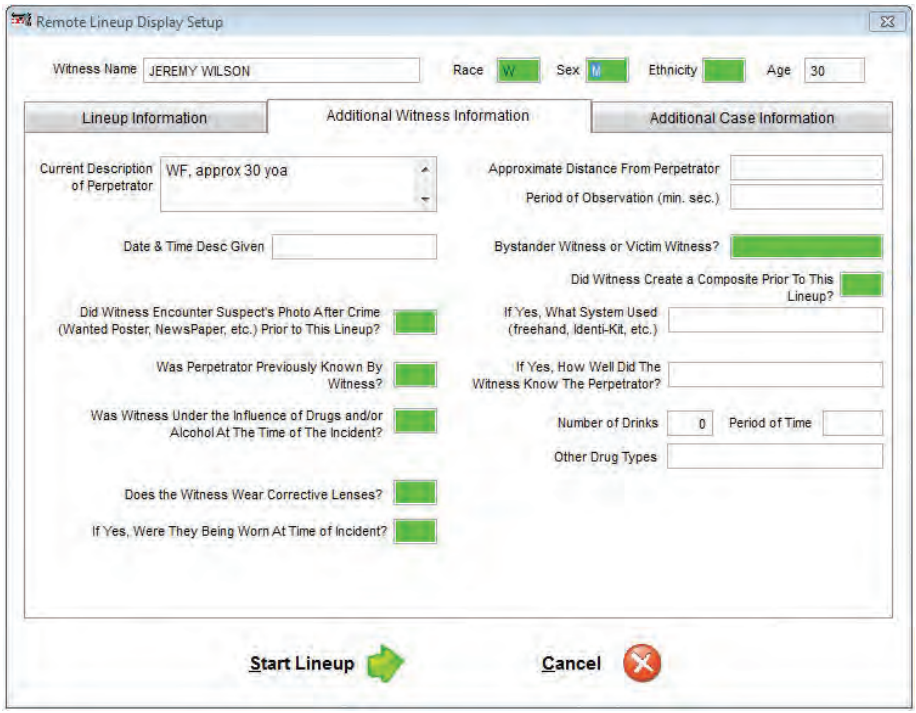
With considerable input from the scientist team and project partners, the Charlotte-Mecklenburg Police

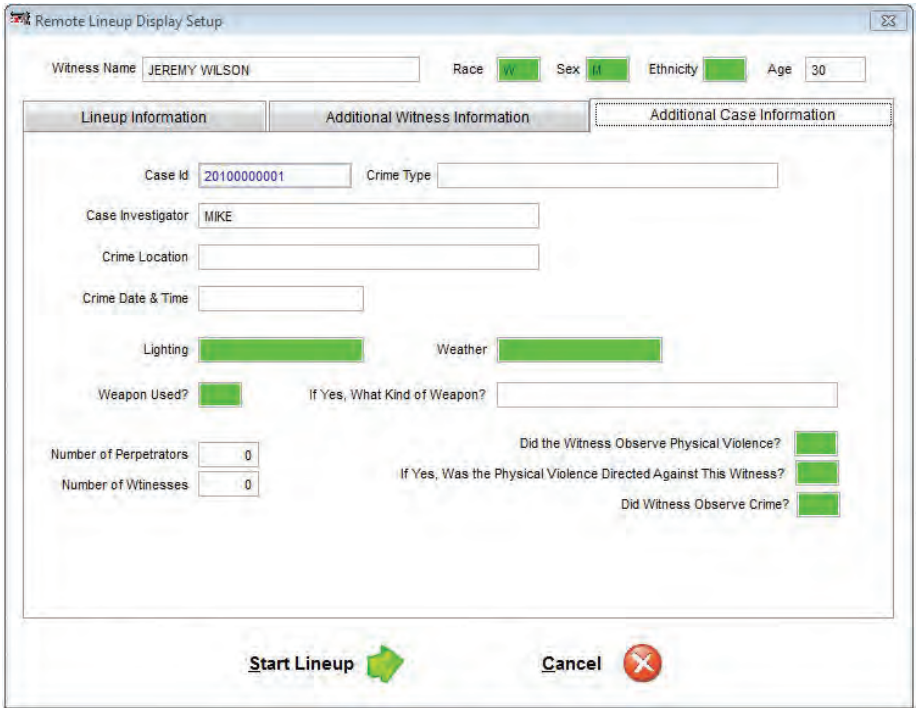
Department and later the other three police departments (Tucson, San Diego, and Austin), SunGard Public Sector, Inc. created a version of the software application that integrated the Greensboro Protocols and captured additional information about the crime (i.e. lighting conditions, witness type, whether the witness knew the perpetrator, etc.) that would aid the scientist team in further analyses beyond the primary research question. In addition, the software application would administer (photo) lineups to eyewitnesses with minimal need for help from a lineup administrator. For the purposes of the study, the case detectives continued to select the filler photos to be used in each lineup but once it was time to administer the lineup to the eyewitness, a second detective did the actual administration using the laptop computer. The software randomly scrambled the photos at the last second (after it was turned over to the witness) and randomly assigned the procedure to be simultaneous or sequential (also only after the computer was turned over to the witness). The computer provided the witnesses with instructions on the procedure (both in writing on the screen and orally with a female voice). The double-blind administrator continued to play a role if the witness had questions about the computer program, and this

17. Mike Garner of SunGard Public Sector, Inc. did excellent work during the course of this project. His time was not charged to the grants received for this project; instead, SunGard Public Sector, Inc. contributed his time and talents. During the project, Garner worked tirelessly to address technical difficulties that arose in creating unique interfaces with mainframe servers that varied from site to site. In some cases, technical difficulties led to delays in software implementation and data collection, despite the best efforts of Garner and information technology professionals in police departments. Law enforcement decisions, unrelated to this field study, to switch the source of filler photos from "mugshots" to driver's license photos or other sources of photos also contributed to some delays. Furthermore, some sites, in particular Austin, had already developed their own interface systems and went to extraordinary lengths to conform to our protocol for purposes of the field study. One lesson that the partners took from these technical challenges is the need to develop new software applications embodying best practices that are compatible with a wide range of different law-enforcement server platforms. Adaptable software applications would allow law enforcement agencies to more easily transition to computer-based photo lineups that might eventually be available for use on handheld devices in the field, allowing for photo lineups to replace show-ups.

administrator made written records of the witness's certainty or other comments if the witness made an identification. The laptop made an audio recording of the entire session and preserved the recording as a WAV file. The recording started when the instructions began and stopped when the lineup administrator ended the session.

Detectives were trained to use the software, which was loaded on laptop computers provided by the study team. The software used an interface with each police department's data base source of photos for conducting photo lineups. Detectives used whatever their usual criteria were for selecting filler photos. For each lineup to be presented, detectives used the software program to record information about the witness and the case. Below are screen shots of the information entered by the detectives before building the lineup[18].





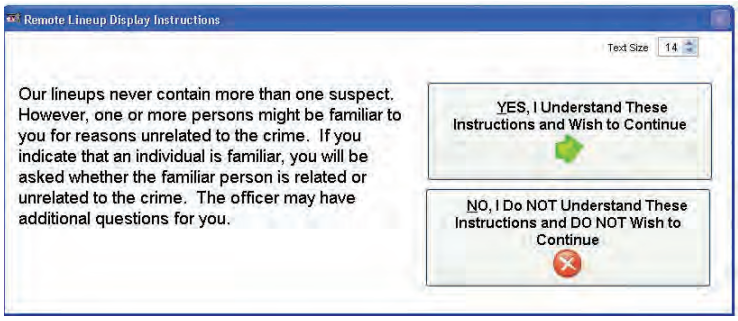
The information filled in by the detectives using these screens became part of a single electronic file for each lineup that was yoked with the lineup photos and all of the identification data. Notice that the information obtained from these screens included "current description of the perpetrator" which will later permit the researchers to assess the extent to which the suspect and fillers fit that description. In addition, information about distance, period of observation, the witness's status as a bystander versus a victim, whether the witness made a composite drawing, whether the witness knew the perpetrator (and if so, how well), information about drugs or alcohol that the witness might have consumed, whether the witness wore glasses, and so on were collected. Detectives would ordinarily be expected to collect this information and document it in some number of police reports, but the software program helped make sure that this information was collected, made part of an electronic record, and was readily available for easy electronic retrieval. The systematic collection of this information permits the researchers to examine the extent to which these factors affect the performance of the witnesses.

Likewise, under the second tab

(Additional Case Information) information was obtained about the type of crime, whether a weapon was involved (and, if so, what kind), whether there was violence involved, the number of witnesses, and the number of perpetrators. Again, detectives would ordinarily collect this information anyway, but the program helped make sure that this information was collected, electronically recorded, and readily retrievable by researchers so they could analyze the extent to which these factors might have affected the performance of witnesses.

Once the lineup was created and the information was entered, the lineup file was uploaded to the police department's server, where it became available for download to any of the laptops that were programmed with the presentation software. The detective then recruited a second person, who did not know which of the lineup members was the suspect and which were fillers, to administer the lineup to the eyewitness. When the administrator opened the designated lineup, the administrator entered additional information, including his or her own name, the name of the eyewitness, the date, time and location where the lineup was shown, and names of any other persons present during the showing. The administrator then began the lineup procedure by cueing up the program and turning the computer over to the eyewitness. Once the witness selected the "**Start Lineup**" button, the computer randomly assigned the lineup to the simultaneous or sequential procedure and randomized the order of the photos, with the caveat that the suspect never appeared in position one[19]. At that point, the computer began giving instructions to the witness.

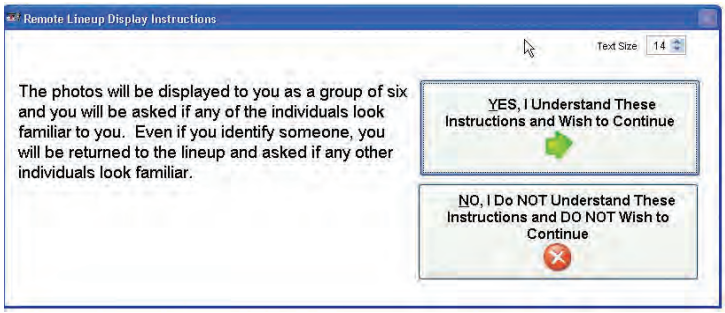
The computer program presented all the instructions to the eyewitness in both written form and via a pre-recorded audio using a female voice. Each instruction was on its own screen and required the witness to acknowledge that she or he understood the instruction before proceeding to the next instruction. These included an early instruction that a lineup contains only one possible suspect; if the witness indicated that someone was familiar she or he would be asked to indicate whether the person was familiar for reasons related to the crime or unrelated to the crime.



This instruction was important and consistent with best practices for two reasons. First, it made it clear to the witness that there was only one suspect in the lineup. Accordingly, if the crime was a multiple-perpetrator offense, the witness would know to not look for any more than one of the perpetrators in any given lineup. Second, because witnesses might see someone in the lineup that they know for other reasons (such as someone from their neighborhood), the instructions made clear that they should indicate that fact so that "**YES**" responses to the familiarity question would yield a record of what they meant by indicating familiarity. Specific witness comments about the familiarity of a lineup member were also captured by audio recording.

If the lineup was simultaneous, the next instruction noted that the photos would be presented as a group of six and that if someone was identified the witness would be returned to the display and asked if any other individuals look familiar. Alternatively, if the lineup was sequential, the next instruction noted that the photos would be shown one at a time and that even if an individual was identified the witness would be shown the remainder of the photos.

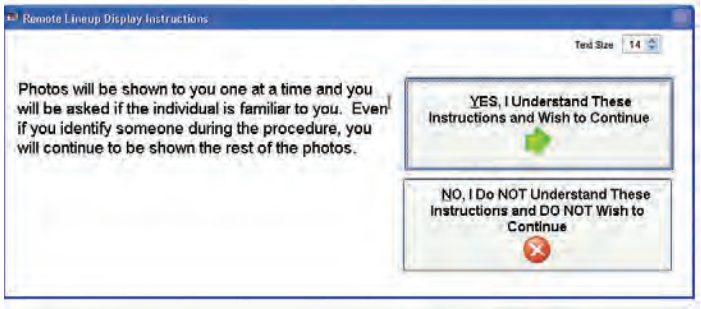
If the lineup was simultaneous:

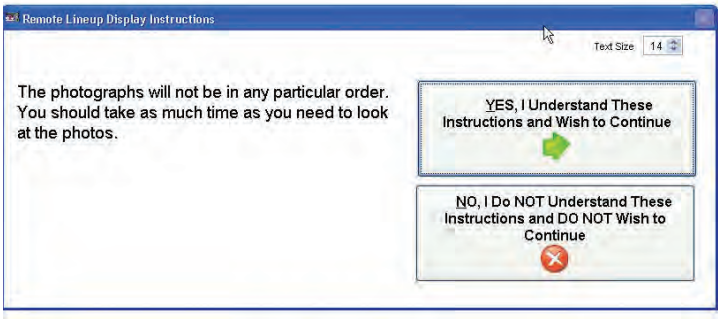


18. Although the content of these screen shots was developed by the eyewitness team of scientists, lawyers, and police who conducted this field experiment, the software underlying the programs is proprietary intellectual property of SunGard Public Sector, Inc.

19. The decision to not place the suspect in position one was to allay any concerns that prosecutors might have about potential defense arguments if the witness identified the first photo they saw in a sequential procedure.

Or, if lineup was sequential:



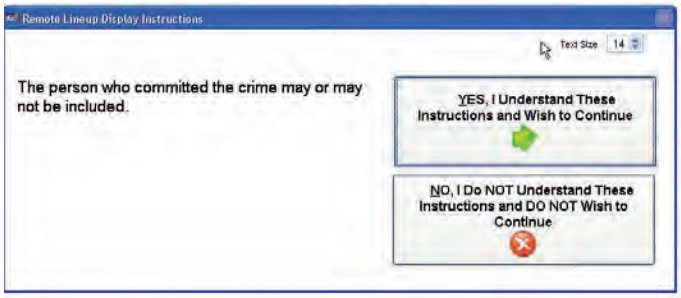


This instruction, pertaining to the photo display, was important to make clear to the witness what form the display of photos would have. This instruction made it clear to witnesses that they would always be returned to the display until they no longer indicated familiarity with any remaining lineup members and that they would view all the photos in the sequential lineup even if they had made an identification. Stating this upfront, before the procedure began, was important so that the witness did not think that returning to the photos after making an identification was some type of "feedback" indicating that a first choice was wrong. Also, the procedure of making sure that the witness viewed all the photos in the sequential procedure (called the continuation feature) was important to avoid cases where the witness might, for example, identify the suspect's photo in position 2 and then not see any more photos. This might create an argument by the defense that the witness was shown only two photos. Furthermore, if the witness identifies a filler early in the sequence, before getting to the suspect's photo, stopping the sequential procedure at that point could prevent the case detective from knowing what the witness might say if she or he were to view the suspect's photo. Even if a witness's identification of a filler undermines the reliability of the witness for purposes of a trial, what the witness says when later getting to the suspect's photo can have investigative value. Other jurisdictions using the sequential procedure implement this "continuation" feature of the sequential lineup.

Other than the instruction pertaining to photo display, all of the instructions were the same for the simultaneous and sequential procedures. The next instruction simply informed the witness that there was no particular order to the photos and to take as much time as needed.

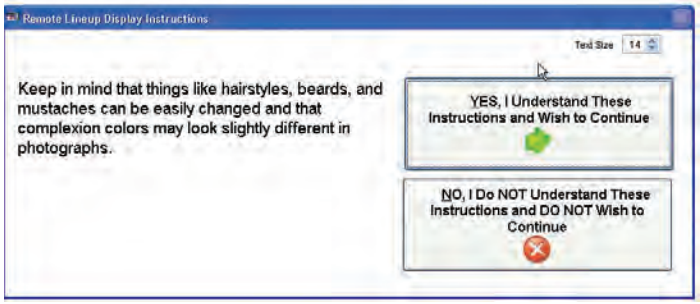
This instruction was important because it helps nullify any implicit theory that some witnesses might have that lineup suspects appear in a particular position in the lineup and it also makes clear that the pace of progression of the lineup is controlled by the witness, not by the computer.

The next instruction was an admonition that the person who committed the crime may or may not have been included in the lineup.



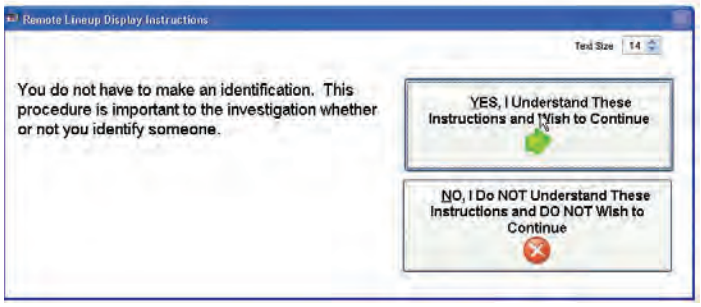
This instruction was critical to the integrity of the lineup procedure. Extensive research has shown that it is essential to disabuse witnesses of the assumption that the perpetrator is in the lineup[20] and this instruction is explicitly recommended by the National Institute of Justice guide for law enforcement on the collection and preservation of eyewitness evidence[21].

The next instruction reminded witnesses that some features, such as facial hair, can be easily changed and that complexion colors may look slightly different in photos.



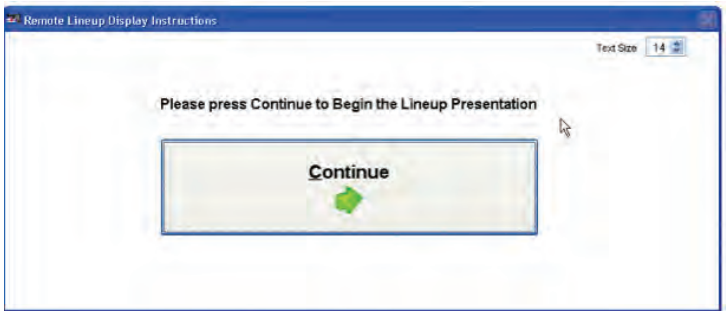
The appearance-change instruction is considered good practice so witnesses do not have the unrealistic expectation that the suspect photo presented by law enforcement for the lineup would necessarily reflect his appearance at the time of the crime. This instruction is also recommended by the NIJ Guide[21].

The next screen told witnesses that they did not have to make an identification and that the investigation would continue even if they did not identify someone.



This instruction helped make sure that the witness would not feel undue pressure to make an identification. Such an instruction is considered a best practice and is recommended by the NIJ Guide (see Footnote 21).

The last instruction screen required the witness to click the continue button to start the lineup.



The next screen was either the six photos of a simultaneous lineup or the first photo of a sequential lineup. Example screen shots for the simultaneous and sequential lineups, respectively, are shown below[22]:

Simultaneous Lineup

20. Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. Journal of Applied Psychology, 66, 482-489; Steblay, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. Law and Human Behavior, 21, 283-298.

21. Technical Working Group for Eyewitness Evidence (1999). *Eyewitness evidence: A guide for law enforcement.* Washington, DC: United States Department of Justice, Office of Justice Programs.

22. "Due to the method used to capture screen images of the software application's lineup instructions and mugshot displays, the quality of the screen images contained within this report are slightly diminished. The software images viewed by witness were very clear, high resolution images."

Sequential Lineup

The images of the lineup members were exactly the same size on the screen regardless of whether the display was simultaneous or sequential[23]. If the witness clicked the "**YES**" button at the top of the screen to indicate that a person or people appeared familiar, the next screen (not shown in this report) asked him or her to confirm whether he or she meant to indicate "**YES**" and gave an option to click "**CONFIRM**" or to click a "go back" button. If the witness confirmed a "**YES**" response for a simultaneous lineup, he or she was asked to click the photo of the familiar person. The clicked photo was then highlighted and the witness again had to confirm that this was the photo he or she intended to select; the witness next indicated whether the person was familiar for reasons related to the crime or unrelated to the crime.

Regardless of whether related or unrelated to the crime, the next screen (not shown in this report) told the witness that the officer had some questions. The lineup administrator, who could also hear the female voice on the computer, then asked the witness to make a statement about the identified person ("How do you know this person?") and, if the witness indicated that this was the person who committed the crime, the lineup administrator asked the witness to use his or her own words to say how sure he or she was that this was the person who committed the crime. Answers to these questions were written down by the lineup administrator and were also part of the audio recorded record.

The software continued to return the witness back to the lineup until the witness indicated that no one else in the lineup was familiar (with the simultaneous procedure) or until the witness had gone through all the lineup members (with the sequential procedure). Witnesses were not told that they could view the sequential lineup a second time. However, if the witness requested a second viewing of the sequential lineup after having gone through the photos, the lineup administrator could initiate a second "lap" of the sequential lineup through a password-secure procedure; the lineup was shown with photos in the same order.

When the lineup ended, the lineup administrator took over the laptop computer and using a yes/no toggle box answered a question about whether any aspect of the protocol could not be followed. If the answer was yes, a text box was provided to explain what aspect could not be followed. The administrator then answered a question about whether he or she (the lineup administrator) knew which person in the lineup was the suspect, again using a yes/no toggle box.

The lineup results were then uploaded to the police department server. A record of all the lineup information (all photos, responses, response latencies, order of photos, whether the lineup was simultaneous or sequential, witness information, case information, and so on) was immediately available as a .pdf document on the laptop for the case detective or others to view the results. In addition, the uploaded file could always be retrieved from the police department server. The audio recording file (a .wav file) was maintained as a separate file that was also uploaded to the police department server. These files could be readily retrieved from the police department server by any of several means, such as via the case number, witness name, or suspect name.

The researchers were provided with these electronic files in the form of both Excel data files as well as the .pdf documents by the police departments via downloading them from the police department server.

23. Keeping the images the same size for the simultaneous and sequential was done only for purposes of this study. One advantage of the sequential lineup in actual practice (rather than this study) is that the sequential images can be larger while still fitting on a screen (or on paper) than they can be when using the simultaneous procedure. However, for purposes of the current study, that would have been a confound for the interpretation of the results and, hence, the image size was kept constant across lineup type.

Creating the Database of Lineups that Followed the Protocol

Across the four sites, a total of 855 lineups were conducted using the laptops and uploaded to the respective department's servers. The numbers of lineups at each site were quite variable for a variety of reasons. The Charlotte-Mecklenburg Police Department, for example, had to discontinue the study soon after it started because of a new law in North Carolina that required all lineups in the state to be done using the sequential method. There were technical problems in Tucson and San Diego with making the software interface with their photo database. The final numbers of lineups were 53 from Charlotte-Mecklenburg, 43 from San Diego, 144 from Tucson, and 615 from Austin, yielding the total of 855 lineups. Among the 855 lineups, 48.8% were sequential and 51.2% were simultaneous. However, 358 of the 855 lineups had to be set aside while testing the primary research questions because of one or more of four problems: 1) the lineup administrator knew which person in the lineup was the suspect, and hence the procedure was not double-blind, 2) the eyewitness knew the suspect at some level of prior familiarity (hence, not a "stranger" identification case), 3) the identification decision of the witness could not be determined (witness picked more than one person and neither the audiotape nor the lineup administrator's notes could disambiguate the question of whether the identification should count as a filler identification or as a suspect identification), or 4) the witness had encountered the suspect or the suspect's photo at some point after the crime and before viewing the lineup. **In other words, the core set of lineups for the central analyses were double-blind lineups from witnesses who were attempting to identify a stranger and who were seeing the suspect's photo for the first time.** The computer documentation from each lineup provided the criteria and the information for decisions about this "protocol-consistent" set of lineups. The following more explicitly describes the criteria for inclusion in the protocol-consistent set:

1. Was it a double-blind lineup? The lineup was considered to be not double-blind if the computer record for any of the following three criteria applied: 1) The lineup administrator answered "**YES**" to the question "Did you know which image was the suspect?" or 2) the case detective and the lineup administrator were the same person or 3) the detective commented (in the record) that the lineup was not performed double-blind.

2. Was it a stranger identification case? The lineup was considered to not be a case of a stranger identification if the detective answered "**YES**" to the question of whether the witness knew the perpetrator.

3. Could it be determined whether an identification was of the suspect versus a filler? There were cases in which a witness identified more than one person. If both were fillers, the witness's decision was considered a filler identification outcome for that lineup. If the witness identified a filler and also the suspect, the researchers - blind to the position of the suspect in the lineup - listened to the audiotape to make a determination as to whether the witness clearly preferred one individual over the other(s) as a "final" decision. If, after reviewing the audiotape, it could not be determined which lineup member was preferred by the witness, it was considered an "unresolved" identification.

4. Had the witness encountered the suspect or the suspect's photo after the crime and before viewing the lineup? The pre-lineup computer program asked the detective if the witness had encountered the suspect's image prior to viewing the lineup (e.g., a picture in the newspaper or on television, or a previous identification attempt such as that from a show-up). If the answer was "**YES**", then the lineup was not included in the protocol-consistent set.

The following is a breakdown of the 358 lineups that were not included for purposes of the current analyses for the following reasons:

Not double-blind = 58.9% of the 358
Not stranger = 34.9% of the 358
Not resolved multiple picks = 6.4% of the 358
Witness encountered suspect or suspect's image prior to lineup = 8.1% of the 358

Notice that the percentages total more than 100% because some of the lineups that failed to fit the study protocol had more than one of these problems and, hence, appear in more than one of the categories.

The lineups that were set aside because they did not meet the protocol for the experiment were equally distributed between the simultaneous and sequential lineups (50.0% were sequential and 50.0% were simultaneous). This is what

would be expected because there is no reason for these protocol breaches to affect how the computer assigned lineups to the simultaneous versus sequential conditions.

Setting aside these 358 lineups that did not meet the protocol, there were 497 lineups that could be analyzed. For purposes of this report, these 497 lineups will be called the "protocol-consistent" set to reflect the fact that they met the protocol standards of being stranger identification cases using double-blind lineup procedures, the suspect or the suspect's image had not been encountered between the time of the crime and the time of the lineup, and the decision of the witness could clearly be categorized as a suspect identification, filler identification, or no identification[24].

The Protocol-Consistent Dataset: Results

Assumptions Tests

Among the first analyses on the protocol-consistent data set were statistical checks on whether some critical assumptions of the study had been met. Particularly important is the assumption of random assignment, both to the simultaneous versus sequential variable and also to the position of the suspect in the lineup. For example, the expectation should be that approximately 50% of the lineups ended up being simultaneous lineups and 50% sequential lineups. In fact, 47.9% were sequential lineups and 52.1% were simultaneous lineups. Neither differs significantly from the expectation of 50%, so these figures are consistent with one of the important assumptions in the experiment.

Because the suspect was never put in position 1 for either the simultaneous or sequential lineup[25], the position of the suspect could be any of five positions, namely 2, 3, 4, 5, or 6. Accordingly, we would expect the suspects to represent about equally often across positions 2-6, or about 20% of the time in each position. Figure 1 shows the actual percentages of time that the lineup's suspect appeared in each of the five possible positions as a function of whether the lineup was simultaneous versus sequential. The percentages of times that the suspect was placed in each position from the random assignment does not differ from that expected by chance for either the simultaneous or the sequential lineups.

Additional checks were made to see if other characteristics of the lineups assigned to the simultaneous versus sequential procedures appeared to be equally distributed. For example, the median number of days between the time of the crime and the time of the lineup for the protocol-consistent data set was 14 days for the sequential and 13 days for the simultaneous. Overall, the assumptions of random assignment seem to have been well met.

Figure 1. Percentages of the Time that the Suspect Appeared in Positions 2 through 6 For the Simultaneous and Sequential Procedures [N = 497]

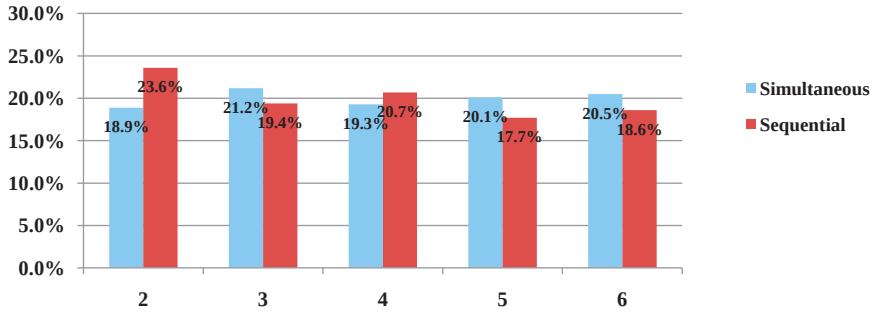
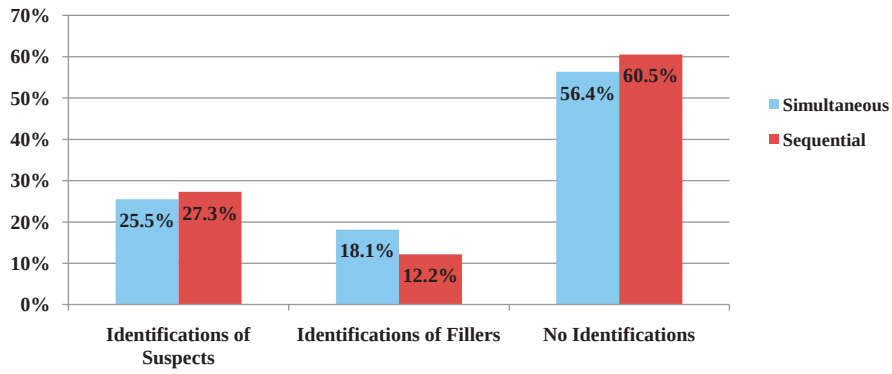


Figure 2. Percentages of Witnesses Identifying the Supect, a Filler, or Making No Identification for the Simultaneous and Sequential Procedures [N = 497]



Main Identification Results

24. The researchers intend to also analyze the non-blind lineups and the non-stranger cases at a later date to see if there are meaningful patterns to be discerned. But these are outside the scope of the questions that were driving this study and are not part of this report.

25. Position 1 in a sequential lineup is the first photo viewed for a sequential linup. For a simultaneous lineup, position 1 is the upper-left corner of a 2 (number of rows) X 3 (number per row) photo-array.

The results of the identification data from the protocol-consistent data set of 497 lineups are shown in Figures 2 and 3. Figure 2 shows that the rates of identifying the suspect were largely the same for the simultaneous versus sequential lineup procedures, with the simultaneous yielding 25.5% suspect identifications and the sequential yielding 27.3% suspect identifications. This small difference was not statistically significant, meaning that the difference is within the margins that could be expected by chance using conventional scientific levels of probability. The rates of filler identifications, however, yielded a larger difference, with 18.1% filler identifications for the simultaneous and 12.2% filler identifications for the sequential. Unlike the suspect identification rates, the filler identification rates produced a difference that is outside of the margins that would be expected by mere chance using conventional scientific levels of probability. More specifically, the probability of obtaining this difference by chance is less than 5%. Hence, using conventional scientific criteria, this difference in filler identification rates is considered to be a reliable difference. Notice as well that the sequential procedure produced fewer identifications overall, a difference that can be accounted for by the lower rate of filler identifications for the sequential procedure.

There are various ways to express these results. For example, it could be noted that the filler identification rate for the sequential procedure is approximately 67% of the rate that was yielded by the simultaneous procedure (i.e., 12.2% ÷ 18.1%) or that the simultaneous procedure produced approximately 1.5 times the rate of filler identifications that the sequential procedure produced (i.e., 18.1% ÷ 12.2%). More meaningful, perhaps, is a calculation that considers both suspect identifications and filler identifications. Figure 3, for example, considers only cases in which the eyewitnesses made identifications and the figure shows the percentages of those identifications that were of the suspect and the percentages that were of a filler. In Figure 3, suspect identification rates and filler identification rates total 100% within the simultaneous and 100% within the sequential because a suspect identification or a filler identification are the only two possible outcomes among those who made an identification. In this analysis, the suspect was identified by 58.4% of those who made an identification using the simultaneous procedure and identified by 69.1% of those who made an identification using the sequential procedure. On the flip side, this means that 41.6% of those who made an identification with the simultaneous procedure identified a filler whereas only 30.9% of those who made an identification using the sequential procedure identified a filler.

"Not Sure" Lineups versus Lineup Rejections

An analysis of "not sure" responses permits an important

Figure 3. Percentages of Suspect and Filler Identifications Among Those Who Attempted an Identification for the Simultaneous and Sequential Procedures [N = 207]

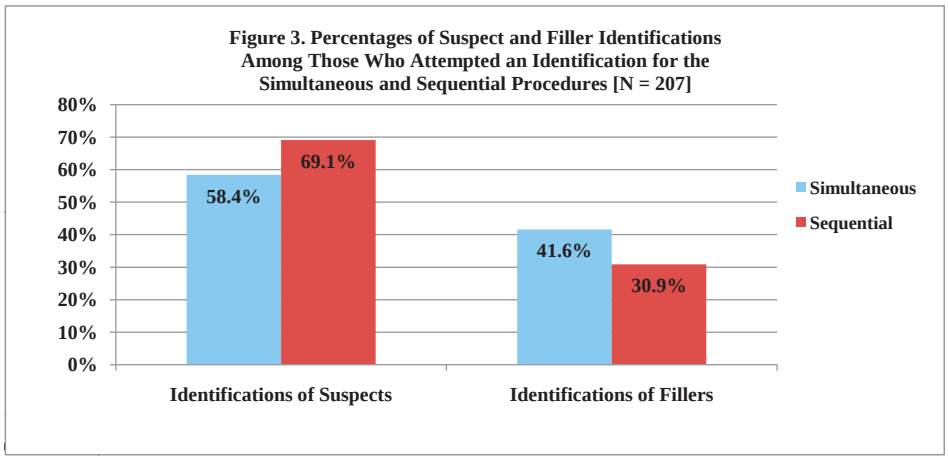
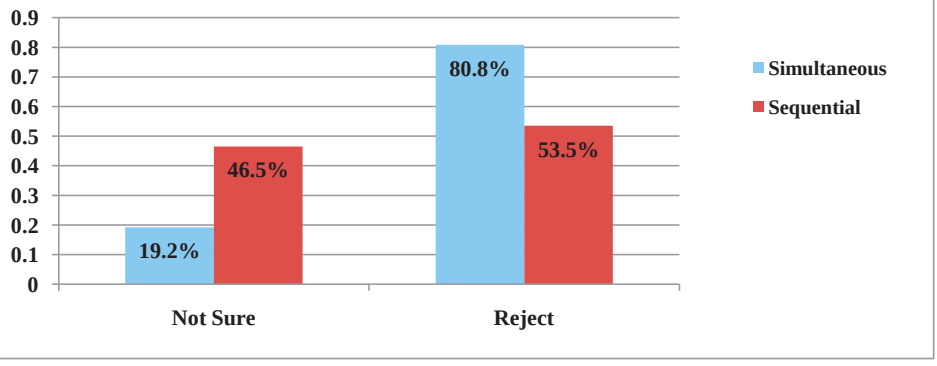


Figure 4. Percentages of "Not Sure" Responses and Lineup Rejection Responses Among Those Making No Identification [N = 288]



distinction to be made between two types of non-identification. One type of non-identification is called a lineup rejection. A lineup rejection is when the witness never says "yes" to any photo *and* never uses the "not sure" option. In other words, a lineup rejection occurs when "no" is the only answer given. A "not sure" lineup, in contrast, is when the witness never says "yes" to any photo but says "not sure" to at least one photo with the sequential procedure or says "not sure" to the set of photos with the simultaneous procedure. Lab research studies have not tended to use a distinction between "not sure" responses and lineup rejections, although there are some notable exceptions[26]. An analysis of "not sure" versus lineup rejections for these field data produced very large differences between the simultaneous and sequential procedures. As shown in Figure 4, for the simultaneous procedure, 80.8% of the non-identifications were of the lineup rejection type and only 19.2% were "not sure" instead of lineup rejections. For the sequential procedure, in contrast, only 53.5% of the non-identifications were of the lineup rejection type and 46.5% were "not sure" responses (Figure 4). Hence, compared to the simultaneous procedure, those using the sequential procedure were not only less likely to identify a filler and just as effective in identifying the suspect, but also less likely to reject the lineup altogether when they did not make an identification. Furthermore, an examination of the "not sure" responses with the sequential procedure found that 28.8% of the "not sure" responses included the suspect. In other words, 28.8% of the witnesses in the sequential conditions who gave a "not sure" response did so to the suspect's photo. This could be an extremely important finding for law enforcement and prosecutors because the "not sure" answer by a witness does not definitively rule out a suspect in a case where there is other evidence, or there are other witnesses, implicating the suspect.

Table 1. Witness Identifications and Non-Identifications as Functions of Suspect Position in Simultaneous and Sequential Lineups [Total N = 497].

		Witness Pick					
Suspect position	No ID	1	2	3	4	5	6
Sequential							
2	66.0%	7.5%	**24.5%**	0.0%	0.0%	1.9%	0.0%
3	56.5%	6.5%	0.0%	**34.8%**	2.2%	0.0%	0.0%
4	69.4%	0.0%	4.1%	2.0%	**20.4%**	0.0%	4.1%
5	63.4%	0.0%	7.3%	0.0%	2.4%	**26.8%**	0.0%
6	55.0%	5.0%	5.0%	0.0%	7.5%	2.5%	**25.0%**
Simultaneous							
2	58.3%	6.3%	**27.1%**	0.0%	4.2%	2.1%	2.1%
3	63.0%	5.6%	3.7%	**20.4%**	0.0%	3.7%	3.7%
4	67.3%	0.0%	4.1%	2.0%	**20.4%**	4.1%	2.0%
5	44.9%	6.1%	4.1%	2.0%	4.1%	**36.7%**	2.0%
6	54.7%	5.7%	7.5%	1.9%	3.8%	1.9%	**24.5%**

Note: Identifications of the suspect are the boldfaced percentages, which run along the diagonal of the table for which suspect position number coincides with witness pick.

Position Effects?

A position effect means that there is a tendency for a witness to be more or less likely to pick a suspect as a function of the where the suspect's photo is in a simultaneous array or the order of the photo in a sequential presentation. One potential concern that has been raised about the sequential lineup is that there might be "position effects." Because the position of the suspect was randomly assigned to position 2-6 for both the simultaneous and sequential lineups, it is possible to look for position effects. Table 1 shows the percentages of time that a witness selected an individual in positions 1-6 as a function of the actual position of the suspect (positions 2-6) for both the sequential and the simultaneous lineup procedures. A position effect is evident to the extent that the percentage of time that the suspect is identified deviates from the expected percentage based on the overall rate of suspect identifications. The boldfaced percentages across the diagonal in Table 1 represent selections of the suspect by position. For the sequential procedure, the suspect was selected between 20.4% of the time (when in position 4) and 34.8% of the time (when in position 3). For the simultaneous procedure, the suspect was selected between 20.4% of the time (in positions 3 and 4) and 36.7% of the time (in position 5). Care should be taken in interpreting these percentages because the sample sizes are small when the data sets are divided into such a large number of categories. For current purposes, however, the important observation is that there is no more evidence of position effects for the sequential than for the simultaneous.

These Field Data Compared to Lab Data

In controlled laboratory experiments, there is no single

26. E.g., Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). On the selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*, 835-844.

"typical" rate that can be used to describe suspect identifications, filler identifications, or lineup rejections. These rates change as a function of numerous variables, including the proportion of lineups for which the actual perpetrator is in the lineup, which is unknown in most field studies. Furthermore, the overall rates of identification and non-identification depend on how good the witnessing conditions are, events that occur between witnessing and the time of the lineup, how well the photograph of the suspects match their actual appearance or their appearance at the time of the crime, and so on. For these reasons, there are likely to be differences between percentages obtained in lab studies and percentages obtained in field studies. We note, for instance, that a recent meta-analysis of controlled laboratory studies showed accurate identification rates to be in the 45%-50% range[27] whereas the current field data produced suspect identification rates of around 27%. The lower suspect identification rate in actual criminal cases could be due to having a modest base rate for suspect guilt (i.e., a fair share of suspects who are innocent). But, it could also be due to poorer witnessing conditions and hence weaker memories by actual witnesses than in the lab studies. Alternatively, the lower suspect identification rates might be due to longer durations between the crime and the lineup in actual cases than in the lab studies or to the use of photos that do not match the appearance of the culprit as well in actual cases as they do in lab studies.

But, it is not the absolute percentages in the lab studies versus the field that are at issue here. Instead it is the pattern of the results that matter. In the context of the simultaneous versus sequential pattern, for instance, the significant reduction in filler identifications that resulted from the sequential lineup procedure in the field experiment is the same pattern observed in the lab studies. There are two reasons to be extremely interested in the reduction of filler identifications. First, filler identifications are the only definitively incorrect response that can be observed in a field study. Identifications of the suspect might or might not be correct, but filler identifications are definitely incorrect. Likewise, making no identification might be a correct decision (the suspect might not be the culprit) or might be an incorrect "miss," but, again, a filler identification is unquestionably a mistake. A second reason to be interested in fewer filler identifications is that filler identifications "spoil" the eyewitness for any later identification attempts should a new suspect surface in the case. For instance, we discovered that in one of the simultaneous lineups the "wrong" suspect was placed in the lineup (someone who shared the name of the suspect). When shown that lineup, the witness picked a filler. That filler identification spoiled the witness in the sense that when the actual suspect was located the witness could not be shown a new lineup without raising serious concerns about the reliability of the eyewitness. Better to get no identification than a filler identification because it keeps the witness unspoiled for a possible new lineup later.

The reduction in filler identifications is especially important in the context of no reduction in identifications of suspects. The main concern that has been raised about the sequential procedure is that it might result in a loss of some accurate identifications even while it reduces mistaken identifications. No evidence supporting that concern was found in these data. This raises an interesting question. Why do lab studies, on average, find that the sequential lineup reduces accurate identifications, albeit to a lesser extent than it reduces mistaken identifications, but the field data do not show a reduction in suspect identifications? The most recent meta-analysis of lab studies of simultaneous versus sequential lineups (Steblay et al, 2011), for example, found an 8% reduction in accurate identifications accompanied by a 22% reduction in mistaken identifications from use of the sequential method.

There are some potentially important differences between the sequential procedure in the current field study and the sequential procedure often used in the lab studies. The current field experiment used a sequential procedure that is more similar to actual practices in the field than is the typical lab procedure. Some lab studies, for instance, stop the sequential procedure as soon as the witness makes a pick whereas the procedure used in this study (and the practice in the field) showed the witness all lineup members even if he or she picks one early in the sequence. The meta-analysis shows that the difference between simultaneous and sequential culprit identification rates in lab studies shrinks to only 5% when the witness is allowed to continue to the end of the lineup. Also, laboratory studies typically use a decision rule for multiple picks in the sequential lineup ("first-choice," or "last-choice") that is an inexact means of determining the

27. Steblay, N., Dysart, J. & 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,* 99-139.

witness's final decision. In this field test, multiple picks were resolved by the witness. The audiotape was reviewed, and the witness's own words were used to determine which lineup member the witness preferred.

Furthermore, the sequential procedure used in the current field study permitted the witness to do a second "lap" if the witness requested it. A second lap increased pick rates by 4.6%. If only the first lap counted, sequential suspect identification rates would have been lower, at 23.5%, which would have been 2% lower than the simultaneous procedure. Moreover, if the lineup were terminated after the first sequential lap, filler identification rates for the sequential would also have been lower, dropping to 10.9% which is 7.2% lower than simultaneous.

It should also be noted that these field data used a somewhat different format than almost all the lab studies. In these field experiments, witnesses always had an explicit "not sure" option available to them and this option was displayed just as prominently as the "yes" and "no" responses to the recognition question. Most lab studies testing the sequential have not explicitly included the "not sure" option. The one lab study that tested this found that an explicit not-sure option reduced witness picks of both culprits and fillers, but led to stronger performance of the sequential lineup; the not-sure option had no effect for simultaneous lineups.

Some social scientists have proposed that the sequential procedure produces a higher decision criterion and this higher decision criterion reduces potential false picks but also raises the chance that a fraction of culprits might not rise above the recognition criterion. Those who consider the simultaneous versus sequential difference merely as a difference in criterion setting will have trouble accommodating these field data, because filler identifications declined but suspect identifications did not. An alternative view is that the difference between the simultaneous and sequential procedures may be seen as a qualitative difference in psychological processes. In fact, the original conceptualization of the simultaneous versus sequential difference was that the simultaneous procedure promoted "relative judgments" involving comparisons between one lineup member and another whereas the sequential procedure promoted comparisons of each lineup member to memory with a more "absolute" decision made about recognition. This original conceptualization tends to predict little or no effect on the witness's ability to identify the perpetrator as long as the perpetrator is present and the witness has a good memory. But, if the perpetrator is not present in the lineup or the witness does not have a good memory, the sequential should reduce mistaken identifications. In this sense the field data tend to fit the qualitative-difference interpretation better than the data fit the criterion-shift interpretation.

Practical Implications of the Results

What do these new field data tell police departments and policy-makers about lineup procedure? Ultimately, that is up to the police departments and policy-makers themselves. But, to the extent that filler identification rates are a reasonable proxy for mistaken identifications of innocent suspects, the sequential procedure should catch fewer innocent suspects in its net. At the same time, there is no evidence from these data that the sequential lineup produces fewer identifications of suspects, at least when the sequential procedure is operationalized the way it was here (double-blind administration of the lineup; witness sees all photos even if an identification is made; second lap permitted if the witness requests it; a clear "not sure" option). Furthermore, there seems to be no practical reason why lineups in actual criminal cases cannot be conducted just as easily using the sequential method as they are using the simultaneous method; there are no meaningful differences between the simultaneous and sequential lineup procedures in effort or time on the part of law enforcement.

At the same time, these field data clearly indicate that the sequential lineup is not a "silver bullet" for the mistaken identification problem. The sequential did better than the simultaneous, but even the sequential procedure still yielded a 31% rate of filler identifications among those who made a selection from the lineup. This is why the eyewitness scientists on this project will look deeper into the data to try to find various "markers" that might help in assessing the reliability of a given identification. For instance, the eyewitness scientists will analyze factors that predict filler identifications versus suspect identifications such as the presence or absence of weapons, whether the witness was a bystander-witness or a victim-witness, the certainty of the witness at the time of identification (extracted from the audiotapes), qualities of the lineup that predict filler and suspect identifications, lighting conditions, duration of the witnessed event, time

passage between the crime and the lineup, whether the witness and perpetrator are of the same or a different race, type of crime, witnesses' verbalizations while viewing the lineup (extracted from the audiotapes), how long it took the witness to make an identification, and numerous other variables. Results of these analyses will come out in later reports and in refereed scientific journal articles.

In addition, subsequent analyses will examine the "non-protocol-consistent" lineups to see how they might differ from the protocol-consistent lineups. For instance, we know little about eyewitness identification performance under conditions in which there was "prior familiarity" between the witness and the suspect. The presumption in the legal system has been that these prior-familiarity situations are much more reliable and in many cases a lineup is never done at all, but instead the witness is shown a single photo to make sure that it is the person that the witness was referring to (e.g., "yes, that is the guy who lives in my building").

Subsequent analyses will also examine the non double-blind lineups to see how their results might differ from the double-blind lineups. Importantly, the decision to conduct the lineup using a non double-blind procedure (e.g., no other detective around to serve as the blind administrator) would be unrelated to whether the lineup was conducted using the simultaneous versus sequential procedure because assignment to the simultaneous versus sequential procedure occurred at the last second (after the witness starts the lineup). And, of course, a computer-generated random assignment would have been indifferent as to whether the lineup administrator was blind or not.

Final Remarks

These are the first data using a double-blind procedure to measure eyewitness identification from lineups for simultaneous versus sequential lineups with actual eyewitnesses. The double-blind aspect of this research is extremely important because it prevents any unintentional influence of the lineup administrator on the eyewitness and thereby takes the lineup administrators' behaviors out of the game as far as interpretations are concerned. This will prove to be particularly valuable when analyzing the certainty data because these are the first field data on eyewitness identification certainty that were collected using double-blind procedures. Of particular interest will be the certainty with which witnesses identify fillers versus suspects and whether this varies by simultaneous versus sequential procedures. Those data are complicated by the fact that witnesses use their own words to describe their certainty (rather than as a number solicited in lab studies). Hence, it will take extra time to have those certainty statements scored using double-blind coders. The current report is only the first "mining" of these data. Later articles will continue to extract additional new findings from this data set.

The method used in this experiment represents the only field test of eyewitnesses using laptop computers to instruct, administer, and record identification decisions from photographic lineups. The software did a great job for purposes of obtaining pristine data, but it turned out to be somewhat clumsy and took longer to use than simply printing the photos and administering the lineup in the traditional way. Because this was the "first generation" of the laptop lineup software, it should not be difficult to make more user-friendly software for the detective and the witness and perhaps make it compatible with a variety of platforms rather than only a PC.

This project was a very successful example of collaboration between numerous groups and individuals. This collaboration involved prosecutors' offices, the Innocence Project, social scientists, the American Judicature Society, the Police Foundation, legal scholars, and law enforcement. [See Acknowledgements]. Also essential to this project were the three foundations that provided the financial backing for this work, without which the project could not have been completed (*Open Society Foundations*, *Laura and John Arnold Foundation*, and the *JEHT Foundation*). This project illustrates the value and potential of collaborations between various entities in addressing an important problem in criminal justice.