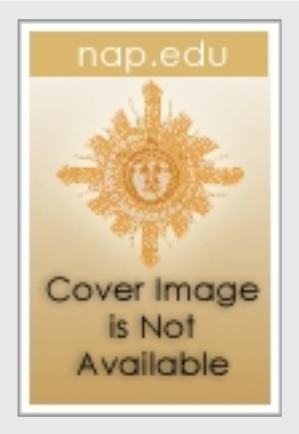

Personnel Selection in the Pattern Evidence Domain of Forensic Science: Proceedings of a Workshop

DETAILS

55 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-45140-6 | DOI 10.17226/23681

AUTHORS

Julie Anne Schuck, Rapporteur; Committee on Workforce Planning Models for Forensic Science: A Workshop; Board on Human-Systems Integration; Division of Behavioral and Social Sciences and Education; National Academies of Sciences, Engineering, and Medicine

Personnel Selection in the Pattern Evidence Domain of Forensic Science: Proceedings of a Workshop

Julie Anne Schuck, *Rapporteur*

Board on Human-Systems Integration

Division of Behavioral and Social Sciences and Education

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

Prepublication copy, uncorrected proofs

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by Contract No. 60NANB150D340 from the U.S. Department of Commerce. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13:
International Standard Book Number-10:
Digital Object Identifier: 10.17226/23681

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. (2016). *Personnel Selection in the Pattern Evidence Domain of Forensic Science: Proceedings of a Workshop.* Washington, DC: The National Academies Press. doi: 10.17226/23681.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C.D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Reports document the evidence-based consensus of an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and committee deliberations. Reports are peer reviewed and are approved by the National Academies of Sciences, Engineering, and Medicine.

Proceedings chronicle the presentations and discussions at a workshop, symposium, or other convening event. The statements and opinions contained in proceedings are those of the participants and have not been endorsed by other participants, the planning committee, or the National Academies of Sciences, Engineering, and Medicine.

For information about other products and activities of the Academies, please visit nationalacademies.org/whatwedo.

COMMITTEE ON WORKFORCE PLANNING MODELS FOR FORENSIC SCIENCE: A WORKSHOP

Frederick L. Oswald (Chair), Rice University

Winfred Arthur, Jr., Texas A&M University

D. Zachary Hambrick, Michigan State University

Andrew S. Imada, A.S. Imada & Associates

Randall S. Murch, Virginia Polytechnic Institute and State University

Ann Marie Ryan, Michigan State University

Jay A. Siegel, Independent Consultant, Michigan State University Emeritus

Nancy T. Tippins, CEB

Julie Anne Schuck, Program Officer

Dixie Gordon, Information Officer

Prepublication copy, uncorrected proofs

BOARD ON HUMAN-SYSTEMS INTEGRATION

Preface

The field of forensic science is often in the news these days and under particular scrutiny. National efforts are underway to develop standards and guidelines for forensic techniques and to push research that tests the accuracy and reproducibility of forensic examinations. Notably, most recently following the workshop, the President's Council of Advisors on Science and Technology published an important report, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*.

In the midst of all these improvements, the allocation and maintenance of organizational, technological, and human resources will continue to play a critical role in the overall performance of a forensics lab. Today, and in the foreseeable future, the quality and reliability of forensic analyses depend in large part on the expertise and capabilities of forensic examiners. In the pattern evidence domain, the field of forensic science is looking to move away from reliance on apprentice-like models of hiring and training toward looking for innovative research-based applications that will effectively identify and test for critical skills and competencies necessary to perform the pattern recognition tasks.

On July 14-15, 2016, the Board of Human-Systems Integration (BOHSI) of the National Academies of Sciences, Engineering, and Medicine sponsored a workshop on personnel selection in forensic science that brought together scholars in industrial and organizational (I-O) psychology, practicing forensic scientists and forensic lab directors, and legal experts. The purpose of the workshop was to develop a better understanding of the current status of selection and training of forensic scientists who specialize in pattern evidence, tools used in I-O psychology to understand elements of a task, and ways that aptitude and performance can be reliably, feasibly, and fairly measured. I thank the National Institute of Standards and Technology (NIST) for the generous support to enable this workshop and publication, particularly Melissa Taylor for having the foresight to bring the I-O psychology, forensic science, and legal disciplines together. In the workshop, we learned a great deal about the task of analyzing pattern evidence and about the state-of-the-art approaches to designing, validating, and evaluating instruments that are used for personnel selection. This summary provides an objective report of what occurred at the workshop, drawing on views presented by individual participants and focusing on the possibilities to develop strategic next steps for a newly formed multidisciplinary community to coordinate collective energies and to continue discussion and improvement efforts.

First and foremost, let me extend my thanks to other workshop steering committee members, who gave generously of their knowledge and time to frame the workshop agenda, identify the presenters, and lend their own expertise to the workshop discussions: Winfred Arthur, Jr., Texas A&M University; D. Zachary Hambrick, Michigan State University; Andrew S. Imada, A.S. Imada & Associates; Randall S. Murch, Virginia Polytechnic Institute and State University; Ann Marie Ryan, Michigan State University; Jay A. Siegel, independent consultant; and Nancy T. Tippins, CEB. The

success of the workshop was critically dependent on a talented and thoughtful group of experts, who took time out of their valuable schedules to present their relevant experiences and research and to participate in what was an interesting and invigorating multidisciplinary discussion: Mark W. Becker, Michigan State University; Wendy S. Becker, Shippensburg University; John M. Collins, Jr., The Forensic Foundations Group; Melissa R. Gische, Federal Bureau of Investigation; Rockne P. Harmon, independent consultant; Scott Highhouse, Bowling Green State University; Bethany Jurs, Transylvania University; Jessica LeCroy, Defense Forensic Science Center; S. Mort McPhail, Society for Industrial and Organizational Psychology; Mara Merlino, Kentucky State University; Liberty Munson, Microsoft Learning Experiences; Daniel Murrie, University of Virginia; Dan Putka, HumRRO; Maria C. Ruggiero, Los Angeles County Sheriff's Department Crime Laboratory; Lisa Scott, University of Florida; and Marvin E. Schechter, defense attorney. In addition, we would particularly like to recognize Susan Ballou, National Institute of Standards and Technology; Wesley Grose, Los Angeles County Sheriff's Department Crime Laboratory; and Heidi Eldridge, RTI International, for their insightful contributions to the discussion, as well as the many others whose interest led them to take time out of their schedules and attend the workshop, either in person or by viewing the webcast.

Also, please let me thank Poornima Madhavan, board director of BOHSI until September 2016, for her ongoing consultation in coordinating and preparing for the workshop. Julie Schuck from the Academies' Committee on Law and Justice deserves special thanks, as she provided invaluable guidance from start to finish: the development of this workshop, assistance in inviting and preparing the participants, coordinating the workshop itself, and serving as the rapporteur for these proceedings, where she thoughtfully captures the many perspectives and major messages presented at the workshop. Dixie Gordon from the Division of Engineering and Physical Sciences and Annalee Gonzales from the Board on Children, Youth, and Families, enlisted to help with administrative aspects of the workshop, kept the workshop on track and running smoothly. My thanks also go to the staff from Sparkstreet for producing the webcast, as well as to other staff within the Division of Behavioral and Social Sciences and Education (DBASSE), especially Viola Horek and Doug Sprunger, who managed communications around this workshop; Eugenia Grohman, who provided consultation with staff on the editing of this summary; Kirsten Sampson Snyder, who managed the report review process; and Yvonne Wise, who managed the production process.

This workshop summary has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Academies of Sciences, Engineering, and Medicine. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the charge. The review comments and draft manuscript remain confidential to protect the integrity of the process.

We thank the following individuals for their review of this report: Heidi Eldridge, Center for Forensic Sciences, RTI International, Research Triangle Park, NC; Henry Swofford, Latent Print Branch, U.S. Army Criminal Investigation Laboratory, Forest

Park, GA; and Wesley P. Grose, Scientific Services Bureau, Los Angeles Sheriff's Department.

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the content of the report, nor did they see the final draft of the report before its release. The review of this report was overseen by John Monahan, University of Virginia. Appointed by the National Academies of Sciences, Engineering, and Medicine, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the author and the institution.

Frederick L. Oswald, Chair
Committee on Workforce Planning Models
for Forensic Science: A Workshop

Contents

1
Introduction

This report summarizes the presentations and discussions at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, held in Washington, DC, in July 2016.[1] The workshop was organized at the request of the National Institute of Standards and Technology (NIST) with the goal of bringing together industrial and organizational (I-O) psychologists, experts on personnel selection and testing, forensic scientists, and other researchers whose work has a nexus with workforce needs in the forensic science field with a focus on pattern evidence.

For the purposes of this workshop, participants focused on the selection and training of forensic scientists who analyze pattern and impression evidence. Such evidence includes patterns produced when an entity comes into contact with a surface or other objects (e.g., fingerprints, shoeprints, toolmarks, and tire treads), as well as patterns and habits considered in handwriting and writing instruments. The ability to detect, interpret, and compare shapes and pattern requires specific visual and cognitive skills.

Workshop participants were asked to review the current status of selection and training of forensic scientists who specialize in pattern evidence and to consider and discuss how tools used in I-O psychology to understand elements of a task and measure aptitude and performance could address challenges in the pattern evidence domain of the forensic sciences.

WORKSHOP OBJECTIVES

The statement of task for the workshop (shown below) guided the appointment of an eight-person steering committee to plan and execute the workshop (see Appendix for list of committee members). At the committee's planning meeting, the focus of the workshop was narrowed to the pattern-evidence domain in forensic science and to the current challenges that exist in selecting forensic examiners from a growing number of individuals interested in the field.

> **Committee Statement of Task:** A steering committee will conduct a 1.5-day workshop to identify and discuss the skillsets for individuals who seek to gain entry into forensic science careers. This will include the identification of human factors variables (cognitive and perceptual skills, decision making abilities, comfort level with technology, etc.), the implications of these variables for education and training of personnel, and their generalizability to the assessment of experienced staff. A summary of the presentations and discussions at the workshop will be prepared by a designated rapporteur in accordance with institutional guidelines.

[1]The archived webcast of this workshop can be found online at http://sites.nationalacademies.org/DBASSE/BOHSI/CurrentProjects/DBASSE_169014#Webcast [October, 2016].

The 1.5-day workshop was held on July 14-15, 2016, with the objective to better understand how the development of selection tools might address the challenges facing the hiring and training of pattern evidence examiners in forensic laboratories. Participants included researchers, industrial and organizational psychologists, forensic scientists and laboratory directors, and others from the criminal justice community. The workshop included 25 presenters and discussants across different disciplines whose remarks helped develop a shared understanding of the task of pattern recognition and the state of the art in personnel selection in the field of industrial and organizational psychology (see Appendix for workshop agenda). An additional 30 people attended, and the workshop was webcast live to reach an audience that spanned the entire country. This proceedings summarizes the presentations and discussion during the workshop.

Remarks made by several participants during the workshop were particularly relevant to the objectives of the workshop and are presented here. The rest of the workshop discussion and presentations is summarized in the following chapters. According to Maria Weir Ruggiero and Wesley Grose (Los Angeles County Sheriff's Department), a relatively small number of people practice forensic science. As such, the challenges faced by forensic laboratories or crime labs[2] are often unnoticed by wider communities. They suggested the workshop and the resulting written proceedings are a chance to make the issues more prominent. Grose recognized that the workshop can help crime labs by providing the opportunity to talk and interact with experts in different disciplines who have different perspectives, skillsets, and knowledge and to find out what other information and resources are available.

Ruggiero noted that in labs and agencies, the personnel and recruitment division is often isolated and separate from the practicing forensic scientists. She said the workshop would be a start toward bridging the gap, providing information that would allow lab personnel to talk about what was needed in terms of examiners' abilities and what could be appropriately measured.

In opening remarks from the workshop sponsor, Melissa Taylor (NIST) outlined NIST's role and vision for the workshop. She noted that NIST has a long history in forensic science, dating back to 1913 with work by Wilmer Souder, and has the necessary research expertise and resources to develop scientific solutions to measurement issues. She hoped the workshop would launch a continuing conversation to answer questions facing the selection of personnel in forensic science, particularly in the pattern evidence domain, where the human is the instrument. Taylor posed a number of questions at the outset of the workshop: How well is the pattern recognition task understood? Are the right people with the right skills in the right roles, and do they have the right information? Do they have the right tools, and the right role models, with the right motivation to do the job? What is needed to be known to fit the best people to the job and the job to the people?

ORGANIZATION OF THE REPORT

This report has been prepared by the workshop rapporteur as a factual summary of what occurred at the workshop. The planning committee's role was limited to planning and convening the workshop. The views contained in the report are those of individual workshop participants and do not necessarily represent the views of all workshop participants, the planning committee, or the National Academies of Sciences, Engineering, and Medicine.

[2]This proceedings uses the terms "forensic laboratories" and "crime labs" interchangeably to recognize the facilities where pattern evidence examiners are employed.

All presenters spoke on their own behalves, often with the disclaimer that their remarks were their own opinions and not those of their affiliated institutions.

This proceedings of the workshop has four chapters with content that generally follows the order of the workshop agenda (see Appendix). However, a decision was made to pull together presentations on the practice and research of pattern recognition in Chapter 2, in order to keep similar topics together and develop a flow for the reader, from a summary of the major issues and insights in this area to discussions on existing tools and potential collaborations. In addition, relevant remarks providing background on I-O psychology and strategies for developing selection tests were integrated as appropriate to keep similar topics and ideas together.

Thus, Chapter 2 focuses on understanding the task of pattern recognition. It opens with description of pattern recognition within forensic science and the current state of education and training for pattern evidence examiners. The second part of the chapter is devoted to different types of research in cognitive psychology that examine the nature of expertise, particularly in regard to visual attention. Research in this area considers individual differences, as well as what attributes may be trainable. Chapter 3 provides an overview of the field of industrial and organizational psychology. It presents some theories and terminology used in the field and lays out the steps in the process of analyzing a job in order to develop and validate tests to be used for the selection of personnel. In Chapter 4, job challenges for forensic examiners beyond the task of pattern recognition are considered (i.e., reporting analyses in the courtroom). The chapter ends with highlights from the workshop discussions and reviews next steps suggested by several presenters that will be useful to continue the conversation between the forensic community and I-O psychologists.

2
The Task of Pattern Recognition

EDUCATIONAL PREPARATION

Before presentations on the current state of training and the nature of the job for pattern evidence examiners, Jay Siegel (Michigan State University, emeritus) provided an overview of changes in the educational preparation for those entering the field of forensic science**.** He pointed out that forensic laboratories have traditionally sought students with a strong science background. In the 1970s, degree programs for forensic science emerged, but there was still a strong preference for science majors, notably in chemistry or biology. In the early 1990s, interest in forensic science exploded, and the number of universities with forensic science programs grew rapidly.

Siegel related that in the early 2000s, the American Academy of Forensic Sciences created the Forensic Science Education Program Accreditation Commission (FEPAC) with the goal of developing rigorous standards for the education of students in forensic science. These standards dictate a curriculum with a solid base of science courses and 15–20 percent of coursework in forensic science. The specific requirements were developed with the recognition that most students would eventually work in the areas of drug analysis, trace analysis, firearms and toolmarks, and forensic biology and that specializations in any other area would require additional curricula or training.[3] Siegel reported about 35 FEPAC-accredited programs in forensic science exist, which offer both bachelor's and master's degrees. Students in accredited programs receive a general coverage of forensic science issues and a review of different types of evidence analyses, according to Siegel, but there is not time to develop and assess specific skills, like pattern recognition, in a college curriculum. Such skill development will likely continue to be part of apprenticeship experiences or on-the-job training. Siegel expressed interest in developing tools beyond the receipt of a degree in forensic science to help identify people who would be successful as pattern evidence examiners.

NATURE OF THE JOB IN FORENSIC LABORATORIES

Jessica LeCroy (Defense Forensic Science Center) presented an overview of what pattern evidence examiners do on a daily basis, their job demands, and the skills required. She contrasted a formal definition of pattern recognition—"a cognitive process that matches information from a stimulus with information retrieved from memory"—with the job where examiners look for patterns and geometric shapes, commit them to memory, and then try to find the same pattern in another source. In the forensic community, pattern recognition is conducted on certain types of evidence (e.g., fingerprints, footwear, tire tracks, ballistics, handwriting, and toolmarks). The task is to simply compare patterns or details from an unknown sample to a known source.

[3]The standards are at http://www.fepac-edu.org/sites/default/files/FEPAC%20Standards%2002192015_1.pdf [July, 2016].

LeCroy showed the audience pictures of fingerprint, tire tracks, and footwear samples. She noted that examiners are expected to look for details and define which areas are suitable to perform a comparison. In the fingerprint sample she showed, only one area could be used for comparison. Examiners use the information or pattern in the evidence from an unknown source to compare to a known source, which may come from a named individual or a database, to determine if there is sufficient similarity to make an association between sources. With fingerprint samples, according to LeCroy, examiners study ridges in the sample, looking at the details, shapes, and voids present in the patterns. Examiners have specialized equipment to aid in this task, which can include magnification tools and digital imaging systems to enlarge and enhance the details present.

Winfred Arthur, Jr. (Texas A&M University) asked why the pattern matching process is not automated and why a human is needed. LeCroy pointed out that some of the process can be automated, notably with fingerprints. An examiner can code details of an unknown and run a search within a database of standards. Some of the challenges with automation include incomplete databases (i.e., there are no standards entered that would match) and search runs that result in "close but no match" outputs. LeCroy asserted that a human is needed to do the coding and decipher any close results. She said an examiner would consider, "Is there sufficient similarity that …additional comparisons [should be conducted] or is there sufficient dissimilarity that [source discounted]?" Siegel added that evidence submitted to forensic laboratories is often not in good condition: for example, fingerprints samples are smudged or a bullet hit a rock after it struck a victim. There can be characteristics on the imperfect evidence that can be analyzed, but it is difficult to tell a machine to look for sufficient details for comparisons.

Because visually detecting shapes and patterns is important and sometimes examiners can look at samples for days to find them, LeCroy said critical attributes of examiners include the cognitive skills necessary to learn, retain, and recall information and the ability to focus for long periods. In her experience, she said, on-the-job training that requires a series of comparisons for long periods and casework with senior examiners can develop and strengthen these skills; however, starting with a foundation of skill and ability, notably visual acuity, makes the training process more efficient and successful. Melissa Gische (Federal Bureau of Investigation) agreed and has found in her role in the laboratory that the ability to see different patterns and detect slight differences in images is not something that can be taught easily.

LeCroy also pointed out the importance of memory to examiners. With tire tracks or shoe impressions, recall of a similar sample in another case can provide a starting point to search for the source. With fingerprints, the examiner is comparing an unknown latent print against multiple records; the process can be more efficient if the examiner can remember specific clusters of information and not constantly refer back as each new record is considered.

John Collins (The Forensic Foundations Group) identified a set of core competences for forensic examiners. His characterization included attributes such as clarity of vision and ability to discern patterns, quality of decision making, ability to exercise self-restraint (e.g., avoiding assigning greater significance to something that is not justified), internal fortitude (e.g., ability to take one's ego out of a decision as well as disagree with others when appropriate), and communication skills. Collins noted that examiners' findings from their analyses need to be communicated to end users (practitioners, attorneys, and judges in the legal system). They should be communicated in a way that does not create confusion or ambiguity and does not assign more or less weight to the evidence than is justified. According to Collins, forensic examiners both perform scientific testing and serve as consultants on the results, and he thought consulting

abilities are currently underdeveloped. The task of providing expert testimony is discussed further in Chapter 4.

LeCroy pointed out that forensic laboratories currently do not have a mechanism to detect existing knowledge, skills, and abilities. Applicants are selected based on other criteria such as minimal educational requirements or relevant experience. The result observed by both LeCroy and Gische is that some of those hired successfully proceed through training and do well as examiners and some do not. Consequences are additional training costs and loss in productivity.

LeCroy said she would like to see tools developed to help measure the skills and abilities necessary to succeed as a pattern evidence examiner. She would also like to see policies and procedures developed and implemented that allow laboratories to use these tools in the hiring process. Gische agreed that having such tools would allow determinations on qualifications earlier in the hiring process before investments in training individuals for 18 months or longer are made. According to LeCroy, the forensic science community needs help validating the importance of visual acuity and related cognitive abilities to the job, determining whether existing tests (like the form blindness test discussed below) are reliable at measuring these skills, and determining the extent training can develop necessary skills and abilities.

HIRING AND TRAINING OF PATTERN EVIDENCE EXAMINERS

Maria Weir Ruggiero (Los Angeles County Sheriff's Department) shared her experiences with hiring and training examiners and talked about recent changes in her department's recruitment and selection process. She reported that the minimal requirements for education or experience are either an associate's degree in any of the specialized areas of photography, crime scene investigation, fingerprinting, and criminalistics or on-the-job experience in an agency working on crime scene investigations, fingerprint comparisons, or as a laboratory technician on automated fingerprint identification systems.[4] The position of forensic identification specialist in the Los Angeles County Sheriff's Department requires an incumbent to conduct crime scene investigations as well as friction ridge (fingerprint) comparisons.

Hiring Examiners

In the past, according to Ruggiero, hiring practices included an appraisal of education and experience to confirm minimum requirements were met, followed by an oral interview that weighted 100 percent toward the hiring decision. In the department's most recent recruitment cycle, Ruggiero reported the addition of a written exam to the process. The written exam covers reading comprehension, written expression, data analysis and interpretation, error analysis, and pattern recognition. Ruggiero focused on the pattern recognition component of the exam. She showed the audience samples of questions where the applicants were asked to identify the image most unlike the other three images in a set. She illustrated that these comparisons were challenging; it was not easy to see the subtle differences. Box 2-1 outlines the number of applications received and how many people met the minimum requirements and went on to pass the written exam. Those that passed the written exam were invited to an oral interview, and top candidates were invited to a second selection interview.

[4]Although the Los Angeles County Sheriff's Department does not require a bachelor's degree for the forensic identification specialist position, Ruggiero reported that 15 of the last 20 people hired for the position had a bachelor's degree in science.

BOX 2-1
Recruitment Case: Los Angeles County Sheriff's Department
December 29, 2015, to January 13, 2016
Approved to fill 12 open positions
Over 450 applications received for the Forensic Identification Specialist position
79 people met minimum requirements and invited to take the written exam
52 people proceeded to the oral interview
38 people successfully passed the oral interview and placed on an eligibility list
13 people placed in the top band of the eligibility list and invited to a selection interview

SOURCE: Ruggiero, M.W. (2016, July 14). Presentation at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, Washington, D.C.
[End Box]

Training Examiners

According to Ruggiero, the latent print comparison training program in the Los Angeles County Sheriff's Department takes approximately 12 months to complete. The program consists of lectures, demonstrations, required reading of relevant literature and technical manuals, and supervised practicum. The bulk of the program is the practicum where the trainees are required to successfully complete a minimum of 600 latent print comparison identifications or exclusions. After successful completion of the training program, Ruggiero reported that the trainees undergo competency testing. The competency testing includes an analytical or technical component and a theoretical component. The technical component consists of eight previously performed comparison cases. To pass, trainees need to correctly identify or exclude all identifiable prints with no erroneous identifications or exclusions. The theoretical component includes a written test and a moot court session. To pass, trainees need to obtain a minimum score of 80 percent on the test and a satisfactory rating on the court session.

Ruggiero recalled that since 2009, the department has hired 23 examiners. Twenty of those hired successfully completed comparison training. In 2009, their lab started administering a form blindness test (see Box 2-2) on the first day of comparison training. Ruggiero observed that over the course of recording scores from the test and monitoring progress in training, the results from the form blindness testing have indicated which trainees would likely excel at training, which would need more attention from the training staff, and which would struggle to complete the training program. Ruggiero said she was not surprised by this observation since, as discussed above, "the ability to perform friction ridge comparisons is dependent upon the examiners' ability to recognize [and interpret] minute differences and similarities."

BOX 2-2
Form Blindness Test

Medically, form blindness is the "inability to see minute differences in form regarding shapes, curves, angles, and size."[a] The earliest reference to form blindness as it applies to forensic work appears to have been written by questioned document examiner Albert Osborn in 1910, who developed a test when a judge was unable to see differences in handwriting samples. He developed a second test more specific to shapes and less specific to handwriting.[b] Many years later, Byrd and Bertram conducted studies of individuals' results on the form blindness test and their ability to complete a training program in fingerprint examination. Their studies included other variations of pretests on pattern recognition. The research has shown correlations between scores on pretests and performance on fingerprint examinations after training.[c]

SOURCE: Ruggiero, M.W. (2016, July 14). Presentation at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, Washington, D.C.

[a]Leisman, G. (1976). *Basic Visual Processes and Learning Disability*. Springfield, IL: Charles C. Thomas.

[b]Osborn, A.S. (1949). *Questioned Documents*, 1st ed. New York: Boyd Printing Co.

[c]Byrd, J.S., and Bertram, D.J. (2003). Form-blindness. *Journal of Forensic Identification, 53*(3), 315-341; Bertram, D.J., Carlan, P.E., Byrd, J.S., and White, J.L. (2010). Screening potential latent fingerprint examiner trainees: The viability of form blindness testing. *Journal of Forensic Identification, 60*(4), 460-476.

[End Box]

RESEARCH ON PATTERN RECOGNITION AND DEVELOPING EXPERTISE

D. Zachary Hambrick (Michigan State University) posited that the core competency under consideration in this workshop discussion is visual attention and the ability to form and maintain a mental representation, even in the face of distractions. He acknowledged that there is a large literature in cognitive psychology on visual attention. Cognitive psychology, Hambrick explained, is the scientific study of memory and thought processes using behavioral and neural methods of discovery in order to better understand individuals' intrinsic skills. At the workshop, four different researchers presented findings from their research on human expertise. Lisa Scott (University of Florida) talked about training perceptual expertise and whether that training was transferable to other sets of stimuli. Bethany Jurs (Transylvania University) discussed how training for latent print examination helps develop the expertise necessary to see through distraction. Mara Merlino (Kentucky State University) reviewed the factors involved in forensic document examination. Mark Becker (Michigan State University) presented what is known about visual search for low-prevalence targets.

Training Perceptual Expertise

According to Scott, visual perceptual expertise is a critical component of expertise required of a number of professions and activities: for example, a Transportation Security Administration agent screening luggage for potential weapons, a radiologist looking for evidence of breast cancer, a geospatial analyst scanning satellite imagery and assessing damage after a

hurricane, and a forensic examiner who conducts fingerprint analyses, as well birding enthusiasts' search for rare species. Scott said she mentioned the latter because her research group often use bird stimuli in their studies.

Scott showed the audience a picture of birds to illustrate how one's entry point of recognition is different depending on one's level of expertise. She asked, "When you see an image, what is the first thing that comes to mind, what is the first label that you come up with?" A novice would look at the picture and identify the animals simply as birds. Someone with a little more experience or who lived close to the ocean might be able to identify the birds as a cormorant, a gull, and a pelican. An expert birder might recognize the birds as a double-breasted cormorant, a western gull, and a brown pelican. According to Scott, each advanced entry point requires the individual to process information at a higher level of specificity.

Scott pointed out that visual perceptual expertise has been investigated across a variety of domains in cognitive psychology and is recognized as an important attribute in how people process information. However, she noted that there is debate on the most important factors for training. Her research examines training factors that might increase perceptual expertise. Her training studies focus on trying to shift people from a basic level of processing, such as basic bird recognition, to the more specific pelican or brown pelican, and examine the best way to accomplish this. She highlighted the effects that were replicated in multiple studies.

Scott pointed out that adults typically process objects and faces at different levels (see Figure 2-1). In her research, she compares training at the basic level to training at the subordinate level. For the study she described at the workshop, the training was a computerized naming species task. An artificial species stimuli was created to help ensure that the effects after training were due to the training and not because of some prior knowledge participants had before training. Scott reported that the training took place training over a 2-week period, and before training and after training, behavioral responses, electroencephalogram (EEG), and eye-tracking were recorded.

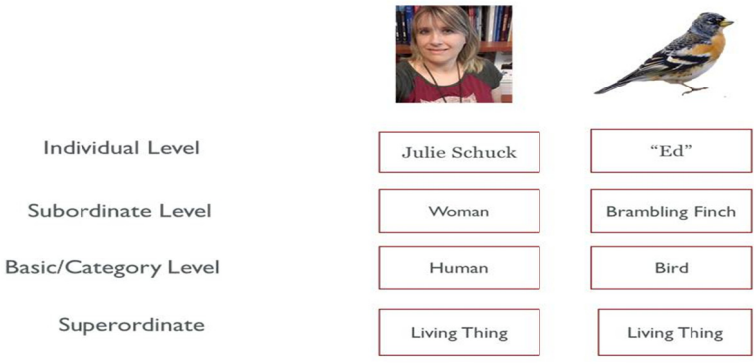

FIGURE 2-1 Perceptual expertise at different levels.
SOURCE: Scott, L. (2016, July 15). Presentation at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, Washington, DC.

Scott explained that in the basic-level training task, participants were shown images of a species and had to distinguish whether the image was part of a "target" family or "another"

family (i.e., yes/no response). For subordinate-level training, participants where shown images of a species and had to distinguish between a set of families and classify the image as one of the families (i.e., labeling response of 1, 2, 3, etc.). For each image response, the participant was given feedback on the correct answer. Scott reported that during the 2-week period, 6 training sessions were held with 25 blocks per session and 900 trials per session.[5]

Her findings suggest that subordinate-level training was important for increasing perceptual expertise from pre-test to post-test. The study also found that training was robust to image manipulations (i.e., performance on degraded images did not decrease significantly) and that training gains generalized from trained exemplars to new exemplars within a family. Scott did not find that such training gains generalized to other families of artificial objects.

Scott informed the audience that EEG records are an added piece of evidence but not typically conclusive on their own. They are useful to show when things are happening in the brain and how fast one is processing information, on the order of milliseconds. In her study, EEG measures confirmed what was found behaviorally by showing significant difference in neural activity after training between stimuli trained at the subordinate compared to the basic level.

For future research, Scott expressed interest in understanding the effects of stress on perceptual expertise after one is trained in a relatively unstressful training paradigm. She asked, "Are there certain individuals who might be more resilient to those kinds of pressures?" She also said she wants to better understand the effect of context on the application of expertise and the interaction between the perceptual information and contextual information. Returning to the birding example, she recognized that birders are not only learning labels, but also learning the context like habitats and sounds made.

Developing Expertise to See through Distractions

Jurs recognized latent print examination and other types of impression analysis as examples of applied visual attention and perceptual expertise. She noted that fingerprint examiners conduct the task of comparing fingerprints in visually demanding environments. The fingerprints are often degraded and there is lots of information, but only some is useful for comparisons; yet they are still able to do this task. Jurs quoted John Vanderkolk from the Indiana State Crime Lab; when asked how examiners do this, he said, "You just have to learn to see through the noise."

Jurs differentiated between two different types of "noise." One is internal noise, which refers to random jitteriness that exists within people and may produce changes in one's decision criteria. The other is external noise, which refers to actual degradation of the fingerprint.

Jurs discussed a training study that investigated how fingerprint examiners' visual systems change across the course of their training to allow them to do this very visually demanding task. The study had a training group and a control group. Both groups had no previous forensic experience and were tested across three different days about a week apart to establish a pre-, mid-, and post-measurement. For each test day, according to Jurs, they were

[5]Scott noted that the trials per block varied based on how many species/family participants had to learn. There were more trials in a 5 species/family block compared to a 1 species/family block. Difficulty was manipulated by changing the progression of how many species they had to identify from each family during each training session. Increasing difficulty started with 1 species/family and worked up to 5 species/family. Decreasing difficulty started with 5 species/family and worked down to 1 species/family.

given three experiments, and in each test, they had the same three experiments.[6] The training group received instruction similar to the first three weeks of fingerprint training at the Indiana State Crime Lab. During the course of training, the technical matching task became progressively harder with fingerprint samples embedded in more external noise.

For the test, Jurs explained, participants were asked to match fingerprints in two different conditions. A fingerprint would come on the screen, either in noise or non-noise, and then four potential matches would come up, and the task was to identify which of the four matched the fingerprint in the middle. "If the individual got the answer right, then the next print that comes on the screen would be lower contrast, so it would be harder to see," she said. "If they got the answer wrong, then the image that came on the screen next would be of higher contrast; it would be easier to see."

Jurs reported for low-noise conditions, participants were very accurate by the third test; there was no difference in performance between the groups, and people maintained very high accuracy. However, for high-noise conditions, the training group outperformed the control group, with significant improvement in the training group's comparison efficiency by the end of training.

Jurs pointed out that even though the control group had the same behavioral improvement as the training group did for low noise conditions, the EEG data showed differences. For the control group, the neural activity between Day 1 and Day 3 increased. For the training group, neural activity between Day 1 and Day 3 decreased. Jurs acknowledged that the significance of increased or decreased neural activity could not be interpreted, but the data illustrate that something different is going on between the groups. She suggested that "training causes the development of perceptual mechanisms that differ from those resulting from just exposure to stimuli."

Jurs reported that participants in both groups had difficulty with the high-noise conditions. However, the training group showed significant improvement by end of training, whereas for the control group, changes were not significant. Even in this short training study, Jurs found effects of training were arguably developing mechanisms that allowed participants to extract out the relevant information and disregard the irrelevant noise.

For future research, Jurs suggested comparing findings from the training group to fingerprint examiners with years of experience. She said she also would track development of other behavioral markers of expertise beyond noise reduction mechanisms across the course of training.

Extracting Information from Handwriting

Merlino discussed research to investigate the way that forensic document examiners extract information out of signature specimens and how they use that information to reach decisions about the genuineness of questioned signatures.

According to Merlino, two key concepts for forensic document examiners are that no two people write exactly alike, which is referred to as inter-writer variability, and nobody writes exactly the same way from one time to the next, known as intra-writer variability. When analyzing signatures, Merlino noted that document examiners recognize the range of variation that can exist and evaluate samples on the consistency of written features, such as the slant of the letters, the writing's orientation to a real or imagined baseline, and the letter spacing. They look

[6]For the purposes of this workshop, Jurs presented findings from two out of the three experiments.

for the presence or absence of the features that they use for the comparison of the questioned and the known documents. They do side-by-side comparisons of the writings that they have. They identify significant differences and similarities, and then they determine whether or not they have a sufficient amount of writing to decide whether or not features of a questioned writing are consistent with the writing habits of the person who produced the known writings.

Merlino presented research that involved eye-tracking technology to record the actual gaze behavior of forensic document examiners as they went about the task of evaluating signatures. Forty-nine fully qualified document examiners and a comparison sample of lay people participated. According to Merlino, the eye-tracking technology records the gaze behavior of participants by tracking the reflection of infrared light from their retinas. This allows the researcher to determine what feature the participant is looking at (fixation); how much total time the participant spends looking at that feature (fixation duration); what features of the questioned writing are being compared to the known writing (referral saccades, or the movement of the gaze from one place to another); the order in which the participant looks at the features (scan path), the total number of fixations on the writing features (fixation count); and the number and total duration of visits to particular areas of the questioned and known writings (visit count and visit duration).[7]

Merlino discussed one protocol where participants were allowed to consider four known signatures for as long as they wanted and then asked to determine whether a questioned signature was a genuine, disguised, or simulated signature; they were not allowed to say inconclusive, but they could provide a value of confidence. Merlino reported that forensic document examiners were statistically significantly better at making these calls than were lay participants across a set of 66 different comparisons. In addition, the research found that the number of years that the document examiner had been in the field was unrelated to his or her accuracy on those calls. Merlino pointed out that forensic document examiners, on average, spent a greater amount of time on their comparisons, and used a greater amount of information than did the lay participants, as indicated by fixations counts, fixation durations, visit counts, and visit durations.

Merlino showed the workshop audience four images of scan paths from the eye-tracking system, one from a lay participant and the other three from professionals. She noted that the lay participant did a very cursory job of looking at the signature and doing the comparison. In the other three images, there was a greater number of fixations and referral saccades. She suggested that the professional examiners sought out more distinguishing characteristics of the signatures and spent more time looking at them than lay participants. She also recognized that even among the qualified examiners, the scan paths indicated a great deal of variation in how examiners go about the visual task.

In another protocol, Merlino reported that single signatures of three different signature types[8] were displayed on the eye-tracking screen for 1 second. Participants were asked to determine whether it was a genuine or a simulated signature. Merlino found that forensic

[7]Merlino explained a person sits in front of an eye-tracking system. Infrared diodes in the system shoot light into the retinas of his or her eyes and the light is reflected off the retina and is recorded by the eye-tracking system to measure gaze behavior. Metrics captured include fixation count, fixation duration, visit count, and visit duration. Merlino pointed out that a visit specified an area on the stimulus and a visit could be multiple fixations within that area of interest.

[8]According to Merlino, text, mixed, and stylized are three different categories of signatures. A text-based signature is a signature where all of the letters or all of the letter forms can be read, as can the allographs within the signature. A stylized signature is essentially a mark that has no distinguishable characteristics to it, and a mixed signature has a combination of text and stylized signatures.

document examiners were able to correctly indicate whether those signatures were genuine or simulated about 70 percent of the time on average. She expressed that the 1-second view was enough for experts to judge such characteristics as internal consistency, line quality, slant, and orientation to baseline. She posited that experts can form images in their minds in a 1-second view and subsequently work with those images.

Merlino and colleagues also gathered information about the training of document examiners by surveying the participating forensic document examiners. Training for the group of document examiners was approximately 2.5 years on average, and subsequent certification was important. Merlino listed favorable aspects of training identified by document examiners: quality materials, textbooks, publications, and actual cases, as well as exercise repetition and immediate feedback.

For future research, Merlino suggested additional studies to examine the facets of the job and better understand the goals of training in order to develop exercises tailored to meet the training needs.

Searching for Low-Prevalence Targets

Mark Becker pointed out that in many important real-world searches, targets are rare. For example, he referred to mammograms. The breast cancer prevalence rate is about 0.3 percent of the scans. Even though radiologists are trained experts, they miss about 30 percent of those cancers. He noted that research shows the prevalence rate (likelihood of a target being in the search) has a large effect on the probability of finding the target. When targets become rare, people tend to miss them more often than when they are common. In research he conducted, Becker found that when targets were present on 90 percent of the search arrays people did relatively well, but when the target was only present on 10 percent of the search arrays, the miss rate went up to 50 percent.

Becker pointed to two sources of errors. One is called the search error, which occurs when people search the display and never get to the target. Becker noted this happens quite frequently and is likely due to a lower quitting threshold. He posited that when targets are rare, people's quitting threshold shifts, and they search less of the display. The second source is an identification error, which occurs when people search and eventually look at a target but fail to identify it as a target. In research models of visual search, according to Becker, this error is "due to a shift in the decision criterion that's being used to evaluate whether each item is a target or not." Becker pointed out that eye-tracking technologies have been used to differentiate between these two types of errors and have shown that the occurrence of both types increase under low-prevalence search conditions.

He and his colleagues have tried various changes to the structure of a search task, while keeping rare targets, in order to identify ways to increase the probability of finding low-prevalence targets. But these efforts have not worked, so research in this area is looking at individual differences given the substantial variability in how well people have performed in these low-prevalence search tasks.

Becker said there is research trying to identify attributes that characterize people who are less likely to commit these errors and therefore are better at low-prevalence search tasks. He said his studies have found that "both types of errors are negatively associated with working memory capacity." In other words, higher working memory capacity reduces the chance of error.

Becker introduced another study that tested for additional predictors of visual search accuracy in these low-prevalence tasks. In this study, measures were collected for four cognitive traits: (1) accuracy on a high-prevalence search, (2) working memory capacity, (3) vigilance (the ability to stay on task), and (4) attentional control. A personality inventory was administered to compare individuals on openness, agreeableness, introversion/extroversion, conscientiousness, and neuroticism.[9] Becker found that all the cognitive traits were significantly correlated with search performance. Of the personality traits, conscientiousness and neuroticism did not correlate with accuracy in low-prevalence search, and openness and agreeableness were not significant predictors. The personality trait of introversion and the four cognitive factors all added significant predictive power to their regression model. Becker reported that their model was able to account for over 50 percent of the variability in low-prevalence search performance. Becker noted that research shows that people likely to perform better in low-prevalence search can be identified by a battery of tasks that are simple and quick to administer. He suggested, more generally, that an individual difference approach to measuring traits or attributes may be an effective way to identify people who would be likely to perform well at tasks that require the accurate detection of rare targets.

The study Becker presented used simplified stimuli in simple displays, in order to control the image statistics. He illustrated a simple search display where people look for a T in any orientation (the target) among other connecting line segments that are not quite a T (e.g., an L or]).

For future research, Becker said he would like to determine if the results he found generalize to more complex stimuli. Additionally, he would like to examine whether individual difference measures could be used to identify the people who would benefit the most from training in visual search tasks.

[9]Mark Becker reported using a 20-question survey, a shortened version of a longer personality inventory, with decent reliability to assess all five personality traits. See Donnellan, M.B, Oswald, F.L, Baird, B.M, and Lucas, R.E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment*, 18(2), 192-203.

3
Industrial and Organizational Psychology

This chapter provides a background on industrial and organizational (I-O) psychology and on strategies for developing selection tests and for recruitment. Remarks from participants and formal presentations at the workshop have been integrated to keep similar topics and ideas together in this chapter of the proceedings.

S. Morton McPhail (Society for Industrial and Organizational Psychology) provided an overview of the field of industrial and organizational psychology to attendees. He cited a definition used on the website of the Society for Industrial and Organizational Psychology: "The scientific study of working and the application of that science to workplace issues facing individuals, teams, and organizations." He recognized that I-O psychology is part of the broader field of psychology but is differentiated by the context in which I-O psychologists study behaviors; they study human behavior in organizations.

Much of the research and work in the field has to do with understanding the nature of human behaviors in the workplace, McPhail explained. Often the goal is to improve aspects of employment such as (1) employers' ability to select and promote qualified people, (2) employees' satisfaction in their work, (3) human effectiveness and productivity, (4) leadership and management, and (5) the workplace in general. McPhail recognized that years of research in I-O psychology have provided knowledge that has brought improvements to many aspects of the workplace. In addition, methods for development and validation of tests have been applied to thousands of different professions. McPhail noted that the field of forensic science can leverage the scientific knowledge and practical experience within I-O psychology to meet current and future needs of the personnel in forensic laboratories.

McPhail pointed to a list of competencies of I-O psychologists taken from guidelines for graduate-level education to illustrate the breadth of the field (see Box 3-1). He noted that three of these competencies are particularly relevant to the workshop: #9 criterion theory and development, #15 job analysis, and #22 personnel recruitment, selection, and placement. Each of these competences represents knowledge of theories and techniques that are used to generate information to improve aspects of employment, as discussed in more detail below. Melissa Taylor (National Institute of Standards and Technology [NIST]) pointed out that all of the competencies listed in Box 3-1, particularly judgment and decision making, would also be useful to the forensic science community. McPhail agreed that many areas of I-O psychology could help. In addition to the consideration of selection tools, the community could get help with measuring and assessing training programs, ensuring reliability of work outcomes, and controlling for bias. These are all things that have been studied in the past in other contexts and could be brought to forensic science, he suggested.

BOX 3-1
Competencies of I-O Psychologists

1. Ethical, Legal, Diversity, and International Issues
2. Fields of Psychology
3. History and Systems of Psychology
4. Professional Skills
5. Research Methods
6. Statistical Methods/Data Analysis
7. Attitude Theory, Measurement, and Change
8. Career Development
9. Criterion Theory and Development
10. Groups and Teams
11. Human Performance
12. Individual Assessment
13. Individual Differences
14. Job Evaluation and Compensation
15. Job/Task/Work Analysis/Competency Modeling and Classification
16. Judgment and Decision-Making
17. Leadership and Management
18. Occupational Health and Safety
19. Organization Development
20. Organization Theory
21. Performance Appraisal/Management
22. Personnel Recruitment, Selection, and Placement
23. Training: Theory, Delivery, Program Design, and Evaluation
24. Work Motivation

SOURCE: Society for Industrial and Organizational Psychology, Inc. (2016). *Guidelines for Education and Training in Industrial/Organizational Psychology*. Bowling Green, OH: Author. Available at http://www.siop.org/ETguidelines.aspx [October, 2016].

<END BOX>

ASSESSMENT FOR THE PURPOSES OF SELECTION

Dan Putka (HumRRO) noted that the benefits of good selection practices and assessments can include improved job performance, reduced turnover, reduced training costs, reduced accidents, reduced counterproductive behavior (e.g., theft, errors), enhanced legal defensibility, and improved applicant perceptions of the employer. Assessment as used in this workshop and in the field of I-O psychology as presented by McPhail implies many different methods and tools, Putka said.

Nancy Tippins (CEB) pointed out that different strategies to identify capable candidates can be used in different contexts and at different points in time, even within the same organization, and that there are a range of different types of selection assessments. For example, résumé reviews and unstructured interviews are considered assessments, as are tests and structured interviews. In some workplaces, one of these strategies may be sufficient. In other

workplaces, several strategies for selection may be useful. For example, inventories of personality traits or tests of cognitive abilities in addition to an interview would provide information useful for selecting the best candidates from a number of applicants.

As discussed further below, any test that is used for selection purposes must be fair, and according to Tippins, citing Title VII of the Civil Rights Act of 1964, "job relevant and consistent with business necessity." There is a well-defined process for developing a fair and appropriate selection test, she noted. The steps of this process include job analysis, test development, criterion development, validation, implementation of the test into recruitment and hiring practices, and technical documentation. Tippins stressed implementation is not an insignificant step and requires concerted attention. She added that documentation is important to record that the test was developed and validated is in accordance with both legal guidelines and professional standards.

JOB ANALYSIS

According to McPhail, job or work analysis comes in many forms, such as task analysis, worker-oriented analysis, and competency analysis. "The work analysis must be sufficiently detailed and complete to identify the key components of accurately defining the requisite knowledge, skills, abilities, and other characteristics, and the performance demands of the work," he said. He explained that a job analysis determines the nature of performing a task, a set of tasks, or a job and includes examining "the physical and social context of the performance and the attributes needed by an incumbent for such performance."

Tippins pointed out that it is essential for I-O psychologists to understand the requirements of a job before developing selection tests or procedures. A job analysis or work analysis is part of the test development and validation process that is designed to identify what the job requires in terms of tasks performed and the knowledge, skills, abilities, and other characteristics necessary to perform these tasks, as well as the environment in which the job is performed and any issues and contingences that affect job performance. Frederick Oswald (Rice University) noted that one reason for understanding the context is the ability to distinguish between issues and requests relevant to selection and those more relevant to training or recruitment.

According to Tippins, before a job analysis is conducted, a project initiation meeting between I-O psychologists and organization leadership is held to confirm the goals of the project, identify any issues or constraints on the research or the intended selection procedures, and review the project plans. Putka acknowledged that it is the role of I-O psychologists to work with organizations and subject matter experts to assemble a list of tasks, identify those tasks critical to perform the job, and use knowledge developed in I-O psychology to identify a set of attributes relevant to those tasks.

In conducting a job analysis, Tippins reported that interviews or focus groups may be held with job incumbents, supervisors, and/or other subject matter experts. Often questionnaires are used to collect quantitative information from people who are very close to the job. These job experts typically are asked to rate the task(s) in terms of frequency of use and importance to job outcomes. They are also asked to rate the selected knowledge, skills, abilities, and other characteristics on importance and the extent to which these attributes are needed upon entry to the job.

Tippins pointed out that "it would not be appropriate to test somebody [in a pre-employment selection program] on a skill or ability that is trained or required at some point after they've been on the job for a while." Findings from a job analysis on critical tasks and necessary attributes of personnel are used to design and choose both the appropriate selection test and the criterion measure (defined below). For some jobs, off-the-shelf tests work fine. For example, she said, there are plenty of arithmetic tests, and one may not need to be developed. If a very specialized skill or area of knowledge is involved, a test designed to assess the skill or knowledge may need to be developed.

THEORIES AND TERMINOLOGY

Presenters defined some terminology commonly used in the field of I-O psychology. The connections between some of these terms are illustrated in Figure 3-1.

Validity

McPhail presented a formal definition of validity from the *Principles for the Validation and Use of Personnel Selection Procedures*[10]: "The degree to which accumulated evidence and theory support specific interpretations of scores from a selection procedure entailed by the proposed uses of that selection procedure."

McPhail emphasized that one does not determine whether a procedure or a test is valid or invalid. It is the interpretations from a procedure (e.g., test scores) that must be shown to be supported by evidence. He listed some conclusions that could be drawn by employers based on candidates' responses or performance on a selection procedure:

- They have (or do not have) the minimum qualifications to perform the work.
- They have (or do not have) the requisite knowledge and skills to competently perform the work.
- They are more (or less) likely to perform the job better than other, competing candidates.

Validity, according to McPhail, is predicated on the evidence supporting the accuracy of these conclusions based on the test outcomes (e.g., scores).

Knowledge, Skills, Abilities, and Other Characteristics (KSAOs)

Tippins noted that KSAOs or KSAPs (for knowledge, skills, abilities, and personal characteristics) are acronyms used often in I-O psychology. Putka explained characteristics can include personality, interests, work values, education, and experience. He pointed to the job analysis as a way of identifying what KSAOs are critical to successful job performance and which ones are irrelevant. It is important to differentiate between KSAOs that can be picked up through on-the-job training or experience and those that are needed upon entry to job (and thus important for selection purposes).

Observed Predictor Measure

Putka explained that an observed predictor measure is a test or assessment used to evaluate attributes that cannot necessarily be seen, such as deductive reasoning or an assessment of conscientiousness. Predictor measures often represent samples or simulations of critical job

[10]Published by the Society for Industrial and Organizational Psychology. Available at http://www.siop.org/_principles/principles.pdf [August, 2016].

tasks that are identified through a job analysis. These measures are developed to test for KSAOs that are critical to performance and needed at entry into a target job, and they are also known as aspects of the latent predictor domain.

Latent Predictor Domain

Putka noted the large body of research in psychology devoted to mapping out individual differences in cognitive, psychomotor, and physical domains. Unobserved individual differences in KSAOs are considered elements of the latent predictor domain.

Observed Criterion Measure

Putka explained an observed criterion measure is an assessment of job performance, testing for outcomes that are actually required on the job. Examples of such measures include (1) ratings from a job incumbent's supervisor or peers and (2) observed performance on a sample of critical work tasks (e.g., samples of fingerprints to match). Putka emphasized that the criterion measure should focus on an individual's behavior and actions and not on job outcomes that would be beyond an individual's control. I-O psychologists develop criterion measures to evaluate and validate selection measures (tests).

Performance Constructs and Latent Criterion Domain

Putka explained that performance constructs are dimensions of performance related to tasks (e.g., analytical and communications skills), context (e.g., teamwork), or counterproductive behaviors (e.g., dishonesty). Unobserved individual differences in performance constructs are considered elements of the latent criterion domain.

Transportability Arguments

In validating selection practices within an organization, Putka explained that transportability arguments could be used in situations where past research or past studies have examined similar jobs with similar predictor measures and similar job performance criteria. Arguments could be made that results from those studies will generalize to the similar situation.

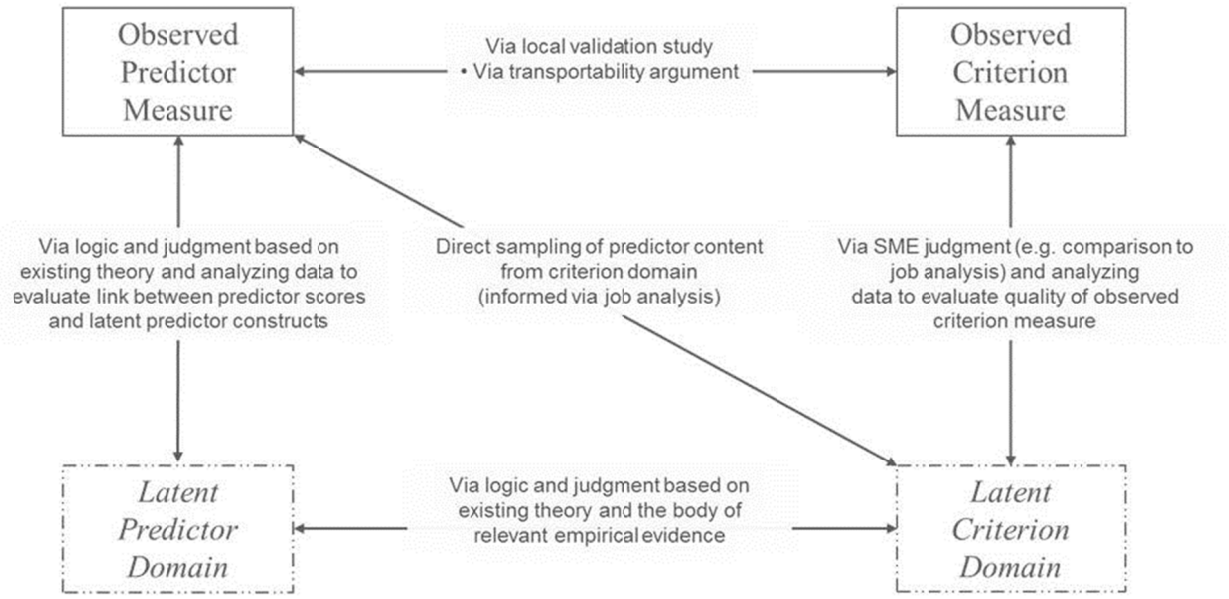

FIGURE 3-1 A framework for validating selection measures.
SOURCE: Putka, D., and Tippins, N. (2016, July 14). Presentation at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, Washington, DC.
NOTE: Dashed lines around the lower boxes indicate domains where characteristics would be latent or unobservable. The lines and corresponding text indicates how the latent domains might be linked to observable measures.

TEST DEVELOPMENT

Liberty Munson (Microsoft Learning Experiences) is a psychometrician who ensures the validity and reliability of tests on regular basis, providing her expertise to 125 tests annually. She oversees the development of certification exams. Such exams serve a different purpose than selection tests;[11] however, Munson pointed out that steps in developing these exams are fundamental to many types of tests. These steps can be broken down into nine development phases: (1) determining the need and rationale for a test, (2) creating a list of skills that could be measured on the test, (3) surveying subject matter experts to determine the importance of these

[11]Munson pointed out that selection tests are used to predict potential performance. In contrast, certification exams are generally used to confirm that someone can actually do and perform the job tasks.

skills and frequency of use, (4) creating and carefully reviewing test items for each skill, (5) field testing the test (beta), (6) revising test items based on feedback from beta test, (7) setting the cut score (passing standard), (8) publishing or implementing the test, and (9) monitoring reliability and validity of the test. According to Munson, different sets of subject matter experts are consulted at each stage of the test development. They provide input for the list of tasks and skills that will be assessed through the job analysis. They usually write test items, and a different set of experts reviews the set of items (questions) for clarity, accuracy, and relevance, she explained. They also take the test as part of field testing and provide comments, and provide guidance when setting the cut score.

During development, Munson noted that results from the field test are analyzed to review the quality of the test items, to determine if they are too difficult or too easy and are able to differentiate between high and low performers. She noted that in field testing, it is important to consider the comments from subject matter experts in addition to the statistical analysis of the results as these can point to issues that do not show up in the statistics. At this point, individual test items may be removed from the exam or revised. Munson emphasized that writing good test items is difficult. Depending on the content domain, subject matter experts or professional test writers create test items; professional test editors typically edit questions created by subject matter experts.

A final step in test development is setting a cut score or standard setting. According to Munson, certification exams generate a pass or fail result with determination of a minimum score that ensures one can do the tasks or skills the exam is designed to measure. Selection tests, on the other hand, are often scored in such a way that allows comparisons among candidates and helps organizations identify the best candidates for the job. Munson explained that a variety of different techniques can be used for standard setting, but many boil down to thinking about the minimal required competence of the target audience and the percentage of questions they should get right. Heidi Eldridge (RTI International), workshop participant, pointed out that on current certification exams for latent fingerprint examiners, a distinction is drawn between knowledge questions where a certain missed percentage is allowable and performance questions where no pattern identifications can be missed.

Munson acknowledged that test development, in general, avoids gate-keeping items (individual test items that must be answered correctly) to allow room for natural error in the process. Oswald confirmed that the reliability of a single test item is typically low. However, multiple test items can form a functional group measure to increase the reliability of assessing a specific skill.

Munson recognized that tests can lose their validity and reliability over time for a number of reasons (e.g., questions are leaked to potential candidates) and emphasized the importance of establishing a review process for the test performance. In some fields, it is challenging to maintain a reliable test as the environment and technology are continually changing. In more stable fields, maintenance is easier, but review still needs to be done. Munson suggested a revisit of the job analysis and test development every 5 to 7 years or more frequently if there are significant changes in the field. In addition, the statistical performance should be monitored as people take the test, and the agency or community should have a process in place to remove items that are no longer valid, reliable, or psychometrically sound and add new content. Test takers can also provide useful feedback via comments and surveys on what they like and do not like, which can be used to improve the test and/or test design and development processes.

VALIDATION PROCESS

According to McPhail, there are many validation strategies for developing and documenting evidential bases. Three have particular prominence in employment testing because they are specifically mentioned in the *Uniform Guidelines on Employee Selection Procedure*[12]: (1) criterion-related; (2) construct-related; and (3) content-related. Putka and Tippins distinguished among the different validation strategies. They noted that all of the strategies involve establishing evidence that scores on a predictor measure are actually predictive of subsequent performance on the job or of some other criterion of interest (e.g., turnover, accidents, production, or counter-productive work behaviors).

In a criterion-related validation study, the evidence is a statistical relationship (e.g., a correlation) that is established between test scores (predictor measure) and the criterion of interest (often job performance measures). Tippins emphasized that the criterion measure needs to be collected in the same way using the same instruments for all people. In the validation of a selection test, job incumbents or applicants can be tested. However, if applicants are tested, often a longer waiting period ensues before accurate measures of performance can be made.

In a construct-related validation, according to Putka, existing theory and relevant literature are used to justify linkages between the predictor measure and criterion. Putka presented the example of Campbell's model of performance determinants.[13] This model depicts a set of distal determinants that help give rise to direct performance determinants. Distal determinants are attributes like a person's ability, personality, interests, work value, education, training, and experience. Although Campbell's model identifies a number of direct performance determinants, the field of I-O psychology has come to consensus on three primary determinants—declarative knowledge, procedural knowledge and skill, and motivation—that have been shown to have a positive relationship with job performance. Putka pointed out that declarative knowledge and procedural knowledge and skill tend to be malleable and therefore amenable to training, whereas attributes like ability and personality are viewed as more stable individual differences that are not as trainable. If a job requires certain ability or personality attributes going in, Putka suggested they should be considered at the selection or recruiting stage and not at the training stage.

In content-related validation, Putka noted that the predictor measure closely resembles what people do on the job. The evidence to link the predictor measure to the criterion is bolstered by judgments of subject matter experts. Tippins explained that subject matter experts are asked to judge the extent different KSAOs are needed to perform the task, as well as to judge the extent to which the test (predictor measure) measures critical KSAOs. She noted that relying on subject matter experts can make the validation process easier in some respects but more difficult in others. According to Tippins, subject matter experts usually know what the job requires in terms of task but may find it tedious or challenging to link KSAOs to the task and test. She noted that sometimes I-O psychologists will use professional test writers instead of job experts to make judgments about the relationships between KSAOs and the tests.

Oswald pointed out that the dichotomy is useful for understanding the differences between strategies, but in practice, organizations interconnect the different validation

[12]Available at http://www.uniformguidelines.com/uniformguidelines.html [October, 2016].

[13]See Campbell, J.P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M.D. Dunnette and L.M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 1), 687–732. Palo Alto, CA: Consulting Psychologists Press.

approaches. Tippins agreed and acknowledged that many organizations strive to continually accumulate evidence of validity. They may start out with a content-related validation study to justify the immediate use of their test. After it is implemented, they might collect applicant and performance data (for those hired) and then conduct a criterion-related study. Inevitably, they need to use the applicant data to adjust for the restriction of range that occurs in the predictor and criterion with the criterion-related study.

PROS AND CONS OF VALIDATION APPROACHES

Tippins highlighted the pros and cons of the different validation approaches and a number of things to consider before deciding how a test is going to be developed and validated for selection purposes. These considerations include: predictive power/validity of interpretations; coverage of job domain; costs; time; staffing environment; personnel requirements; adverse impact and other legal risks; applicant reactions; and type of return on investment study possible. Box 3-2 summarizes considerations for content-related and criterion-related validation strategies.

According to Tippins, a content-related validation study typically takes less time than a criterion-related validation study. Theoretically, both types of studies could be done in about the same amount of time, but in reality, Tippins opined that "it is a lot easier [and faster] to get a smaller group of subject matter experts together in a room to [make judgments] than it is to test several hundred people and collect performance data from their supervisors or their trainers." In addition, content-related validation studies are usually less expensive than criterion-related validation studies. Tippins pointed out that smaller samples can be used in content validation. A sample that is representative of the various demographic groups for the job is needed but it does not have to be as large as for criterion-related validation studies.

However, Tippins noted content-related validation is not useful in all situations. The *Uniform Guidelines* states that content validation approaches are not appropriate for measures of certain constructs, specifically calling out personality and intelligence. Tippins explained that as the construct being measured becomes more abstract or difficult to observe, appropriateness and legal defensibility of content-related approach diminishes. In addition, Tippins conveyed that a content validation study does not provide evidence needed to address fairness concerns and demonstrate that the test works equally well from all members of protected classes, primarily defined by race and gender.

BOX 3-2
Comparison of Validation Approaches

Content-related validation

- Relies on subject matter expert (SME) judgment
- Typically takes less time to complete
- Lower financial investment
- Does not require collection of any criterion
- Can utilize smaller, representative samples
- SMEs are usually willing to participate
- May not be sufficient for measures of personality and intelligence
- Cannot address fairness/bias concerns

Criterion-related validation

- Relies on statistical relationship between predictor and criterion
- Typically takes a longer time to collect predictor and criterion data
- Can be expensive, particularly for concurrent studies
- No good, easy-to-collect criterion measures for performance exist; developing criteria for employees' potential or progression take a long time
- Requires a large, representative sample
- Incumbents often do not like to be tested
- Leaves organizations vulnerable during the validation period if applicants are used instead of incumbents
- May be the "gold standard" for validity evidence in litigation

Source: Putka, D., and Tippins, N. (2016, July 14). Presentation at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, Washington, DC.

<END BOX>

PERSONNEL RECRUITMENT

In her workshop presentation, Ann Marie Ryan (Michigan State University) pointed out the importance of recruitment and how it feeds into selection. Recruitment is the process of getting people in the pipeline ready to have the skills and qualifications to be selected for the job and attracting them to apply. Today, through advances in communications and social media as well as increased networking, more people are exposed to any single job posting compared to a decade or two ago. The downside of this is the heightened possibility of many candidates without the requisite qualifications and fit entering the recruitment process.

Earlier in the workshop, McPhail emphasized that the goal of personnel recruitment should be the effective matching of the needs, preferences, skills, and abilities of job recruits and those of existing employees with the needs and preferences of organizations. Ryan reminded the audience that employers, when trying to get personnel selection right, should keep recruitment in mind and consider whether their recruitment efforts will return enough applicants of the right quality and skills to make informed decisions on hiring. She recognized that some messages

allow prospective applicants to screen themselves out ("self-selection"), which in some situations is good. However, she cautioned against messages that might inadvertently lead those who should not screen themselves out to do so. Employers must consider whether their message is appropriately screening-oriented. Does it make the job look too hard to get? Or does it give people good information so informed decisions can be made about the fit? Who does the message attract? This becomes particularly important in fields that are currently not attracting a demographically diverse group of people.

Ryan distinguished between external and internal recruitment. External recruitment involves hiring people from outside the organization. Internal recruitment refers to applicants from other jobs within the organization. There are benefits to internal recruitment: applicants know something about the organization and the job, and hiring managers can access information about their performance easier. It can also build morale by defining career paths. On the other hand, she said, relying too heavily on internal recruitment can lead to stagnation in creativity or improvements if fresh perspectives are not entering the organization or create significant vacancies in other areas of the organization.

Ryan asked the audience to consider the current situation in hiring pattern evidence examiners in terms of internal and external recruitment. She posed a series of questions: Is it a good balance? Are the pros and cons appropriately considered? Are people who are potential candidates getting appropriate information about what the job requires to make an informed decision to apply or not apply? Ryan pointed to issues of recruitment sources. In today's networking environment, people get information about jobs through social networks and often from multiple sources[14] and employers have less control over their job descriptions and perceptions of their industry on the Internet and in different communities. She noted that private-sector organizations are starting to focus on the role of all their employees in recruitment—what messages are given to employees and what they share in their networks.

Ryan acknowledged that research shows that referrals from current employees are effective toward identifying more successful performers. However, it is also well established that reliance on referrals can result in greater homogeneity in the types of people in an organization. This can be a negative in situations that do not have a very diverse group of employees, and as a result, an organization continues to end up with a very homogeneous group.

In his remarks, Scott Highhouse (Bowling Green State University) focused on what is known about getting the right personnel fit for a workplace. He introduced Schneider's ASA model, which begins with a process of attraction where an applicant selects the workplace with the characteristics the applicant desires, followed by a process of selection wherein the employer selects applicants with the characteristics it desires, and lastly by a process of attrition when people leave the organization if they do not fit within it.[15] Organizations generally want to limit the last stage as much as possible.

Highhouse conveyed that the current literature on attraction[16] suggests that an organization's recruitment information and messages lead an applicant to have an image of the

[14]Ryan reported that 73 percent of 18 to 34-year-olds found their last job through a social network, citing J.P. Medved (2014), Top 15 recruiting statistics for 2014. Available: http://blog.capterra.com/top-15-recruiting-statistics-2014/ [October, 2016].

[15]For a description of the attraction-selection-attrition (ASA) model, see Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*, 437-454.

[16]This literature was recently summarized in a review by Lievens and Slaughter. Lievens, F., and Slaughter, J.E. (2016). Employer image and employer branding: What we know and what we need to know. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*, 407-440.

employer. The field of I-O psychology approaches the employer image in terms of instrumental dimensions and symbolic dimensions. Instrumental dimensions are the tangible things, such as location, pay, benefits, and advancement, that people consider when looking at a job. Symbolic dimensions are subjective perceptions, for example, whether people perceive an employer to be innovative, dominant in the industry, sincere, competent, or other qualities. Highhouse emphasized that symbolic dimensions can be important for attracting applicants if they positively distinguish an organization from other workplaces even when the instrumental dimensions are the same.

In developing image dimensions, Highhouse suggested that employers examine competing organizations and determine how their image might differ from other potential employers, identify the instrumental and symbolic dimensions that distinguish competing workplaces, and then choose to emphasize those dimensions that provide a competitive advantage and attraction.

Highhouse recognized that attraction may not be an issue in forensic science. The issue may be avoiding misfit. A common approach to limit misfit and avoid unrealistic expectations is the realistic job preview. The approach often includes employee testimonials, video presentations, and/or work simulations to illustrate aspects of the job that are both favorable and unfavorable. He noted that areas with the potential to create discomfort for some people can be identified in a job analysis (discussed above) or through critical incident interviews.[17] Part of the selection process entails understanding the discomforts of a job and identifying candidates who would be less aggravated by these discomforts.[18] Highhouse recognized that selecting applicants who are more tolerable of any job discomforts does not preclude redesigning the job to reduce discomforts.

PROFESSIONAL STANDARDS AND LEGAL REQUIREMENTS

As McPhail noted, in the United States, assessment in the employment context is governed by a broad array of legal and regulatory requirements intended to provide fairness, which various state and federal agencies enforce. Box 3-3 shows the statutes and enforcement agencies most likely relevant with respect to assessment and selection.

[17]Critical incident interviews involve asking job incumbents to identify situations on the job that were especially uncomfortable.

[18]At the workshop, participants noted that the job of a pattern evidence examiner often requires staring at patterns and samples for days and occasionally puts people in situations where their judgment is questioned**.** These experiences may be too aggravating for some people, whereas others may find them satisfying or less egregious. See Bernardin, H.J. (1987). Development and validation of a forced choice scale to measure job-related discomfort among customer service representatives. *Academy of Management Journal*, *30*(1), 162-173.

BOX 3-3
Employment-Related Statutes, Guidelines, and Enforcement Agencies

Statutes and Guidelines

- Civil Rights Act of 1964
- Executive Order 11246 (1965)
- Age Discrimination in Employment Act of 1967
- Vocational Rehabilitation Act of 1973
- *Uniform Guidelines on Employee Selection Procedures* (1978)
- Americans with Disabilities Act of 1990
- Civil Rights Act of 1991
- Extensive case law in both state and federal courts

Enforcement Agencies

- Equal Employment Opportunity Commission (EEOC)
- Office of Federal Contract Compliance Programs (OFCCP)
- Department of Justice (DoJ)
- Department of Labor (DoL)
- State human or civil rights agencies

SOURCE: McPhail, S.M. (2016, July 14). Presentation at the Workshop on Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners, Washington, DC.
<END BOX>

The bottom line of all these statutes and guidelines, according to McPhail, is that "employers may not discriminate on the basis of race/ethnicity, color, gender, national origin, religion, age, or disability in *any* aspect of the employment situation." He emphasized that this requirement applies to all selection procedures, and noted that any procedure that has disparate (adverse) impact[19] must be shown to be valid. In the legal sense, "valid" refers to demonstrating that procedures are job-related and consistent with business necessity. He stressed for legal defensibility, the laws that govern selection are primarily associated with enforcement of fairness in employment selection and hiring, but they apply to any aspect of the employment situation, including promotions or transfers. Pre-employment selection was the initiating point and continues to be probably the single most common, he said. The defensibility comes from having done the validation research in a way that is compliant with professional standards, the Uniform Guidelines and case law. For defensibility in a job analysis, he said, "You've sampled adequately, you've conducted the interviews, you've gathered data, you've applied reasonable scientific principles to the obtaining of facts regarding the interpretations you can make of the test scores, [and] the evidence supports that."

Ryan pointed out that the legal requirements are important because they constrain and define some of the things that can be done. However, more importantly, are the professional

[19]Disparate impact means that the proportion of successful candidates from one subgroup is substantially less than the proportion successful from another subgroup.

standards for the I-O psychology community. They dictate finding the best, most accurate ways to predict who is going to be an effective employee.

4
Other Considerations and Next Steps

The workshop sought to understand how the development of fair, valid, and reliable selection tools might address some of the challenges facing the hiring and training of pattern evidence examiners in forensic laboratories. The focus was on the pattern recognition task required of the job and consideration as to whether any individual traits or abilities were necessary before hiring in order to perform successfully on the job and in training. However, the job of a pattern evidence examiner is not just analyzing evidence; it sometimes entails reporting findings to end users, such as lawyers and judges.

UNDERSTANDING EXPERT TESTIMONY

The workshop steering committee included presentations on how expert testimony is treated in the courtroom in order to better understand what preparation pattern evidence examiners need. To do this, the steering committee invited four people to present their perspectives: Rockne Harmon is a retired prosecutor who served in California, Marvin Schechter is a defense attorney in New York, Mara Merlino (Kentucky State University) is a researcher whose research examined expert testimony, and Dan Murrie (University of Virginia) is a forensic psychologist who has trained others to testify.

Harmon laid out the nature of the legal system and presented some questions to consider if the forensic science community embarks on developing selection tools. He pointed out that the legal system is an adversarial system. “By definition, there have to be two inconsistent, irreconcilable goals,” he said. In the courtroom, according to Harmon, a defense attorney reserves the right and has the responsibility to challenge testimony on evidence. He noted that most criticisms are not related to the findings but have more to do with the process, how evidence is collected and handled, and how analyses are documented. He recognized that forensic testimony occurs in context and is seldom the only testimony presented in a case.

Harmon identified several questions that might be worth considering before selection tools are put into practice. The legal system has a process called discovery, which generally requires the government to turn over all information relevant to a case. Harmon queried whether selection tools would be deemed discoverable or whether there would be some privilege that applies. He considered that selection tools could be used to bolster testimony if they are shown to speak to an expert’s qualifications.

Schechter concurred that the right to cross-examine is one of the hallmarks of the legal system. He recognized that the courtroom context places a number of challenges on expert testimony (e.g., a judge may decide to exclude the evidence from the case or put limitations on what can be reported to the jury, and attorneys will question methodologies and interpretations). However, he pointed out the toughest witness to cross-examine is the one who wants to give objective information.

Merlino remarked on the admissibility of expert testimony in court. She pointed out that judges are required to look at the qualifications of experts and the scientific merits of the experts’ methods and conclusions. These determinations are often the result of interactions that take place

among attorneys, experts, and judges and can at times be directed by precedent. Merlino reminded the audience that in this context communication skills become particularly important. A pattern evidence examiner asked to review evidence in the courtroom will be expected to communicate clearly the information judges and attorneys need to make good decisions and effectively represent their clients.

Schechter pointed out that the nature of forensic testimonies and the culture within forensic laboratories have been in the process of changing over the last 5 years or so. He observed these changes are for the better. He emphasized that an examiner's ability to perform and testify successfully in the courtroom will require an understanding of the courtroom context and training to communicate appropriately in this context.

Murrie recognized that forensic examiners need certain skills to provide good testimony in court. They should be able to communicate well, particularly under pressure, and have the right amount of ego to recognize they have specialized knowledge and can present it in a neutral, objective way. Murrie pointed out that such communication is challenging because most people are influenced by their context. "It's really difficult to be an objective scientist in an adversarial system," he stated.

Murrie referred to a 2009 report from the National Research Council[20] that highlighted the vulnerability to certain cognitive and contextual biases and to studies by Itiel Dror that show contextual information (e.g., knowledge of a confession) can influence expert opinions.[21] He emphasized the importance of understanding that bias is not an ethical issue. "Biases are universal and automatic. They happen without awareness," he stated. He recognized that bias can be more problematic in situations where the data are ambiguous (e.g., when analyzing low-quality, degraded evidence). The challenge, according to Murrie, is that errors due to bias cannot be solely avoided by good intentions. He added that general knowledge of bias will not eliminate biased judgments.

Historically, according to Murrie, the forensic science field did not do much to manage bias. However, increasing attention to cognitive bias has prompted a number of changes in the field and interventions. Murrie referred to blinding and case management procedures that separate contextual information from the analysis task. He surmised that progress in managing bias will continue to be made procedurally by changing lab procedures in ways that account for human bias. He cautioned, however, that some procedural changes may have implications for job satisfaction since some examiners appreciate exposure to case details.

On another note, Murrie pointed out little is known about examiners' motivations. He considered that different examiners might have different motivations for the job, such as fighting crime, solving cases, or doing science, and that these motivations might lead toward different orientations toward the task of pattern recognition.

CONTINUING THE DISCUSSION

In closing the workshop, Frederick Oswald (Rice University) observed that the attendees had created a community of interest among different disciplines. They had discussed issues and

[20]National Research Council. (2009). *Strengthening Forensic Science: A Path Forward.* Washington, DC: National Academies Press.

[21]Dror, I.E. and Charlton, D. (2006). Cognitive bias and fingerprint examiners: Why experts make errors. *Journal of Forensic Identification, 56*(4), 600-616. Dror, I.E., McCormack, B.M., and Epstein, J. (2015). Cognitive bias and its impact on expert witnesses and the court. *The Judges' Journal, 54*(4), 8-15.

generated ideas on how to move forward. He suggested concerted efforts should be made to continue the conversation to understand each other's disciplines.

Wendy Becker (Shippensburg University), Andrew Imada (A.S. Imada & Associates), and Ann Marie Ryan (Michigan State University) presented a summary of the workshop discussion and highlighted ways that they said the fields of forensic science and industrial and organizational (I-O) psychology could work together in the future.

Becker started by outlining some of the contextual factors that would continue to impact personnel selection in forensic science. Imada and Ryan then presented some ideas on next steps. Ryan enumerated several contextual challenges that exist in the field of forensic science that will continue to impact personnel selection. One of these challenges is budget. She recognized that workshop presenters had illustrated uneven budgets, where at times there are appropriate resources to hire a new cohort of forensic examiners and at other times hiring is put on hold. There is also the influence of the civil service environment and restrictions that can impose long waits in the hiring process. Several participants recalled lab managers reviewing a list of candidates only to find the best candidates were no longer available.

Becker recognized changes in progress in the forensic science field that will affect the management of forensic laboratories and the culture of these workplaces. These changes include the creation of independent public laboratories, separation of labs from police departments, rise of private laboratories, and expansion in accreditation. Such changes can have positive and negative effects on the hiring process and selection, she said.

In the course of the workshop, Imada said he identified what he called "islands of innovation" in current selection systems. He referred to the form blindness test that had been developed many years ago but is being used today in predictive ways. He said interpretations of this test were consistent with what was heard from the research presentations on cognitive differences in visual attention. He referred to a selection process in a lab in a large city that was developed by professional staff and employed what is considered a funnel approach.[22] Imada felt this example provided a model that a sound selection process could be developed and implemented. He heard about other signs of change within forensic laboratories. He pointed to remarks made at the workshop that forensic examiners generally are beginning to report findings on evidence as "associated with" as opposed to "identified with" known sources or suspects.

Imada recognized that recruitment is not a problem in forensic science. The problem is more likely selection ratio. Attendees heard that in Los Angeles this year, 450 people applied for 12 positions (see Chapter 2). Although Imada felt this was a good problem to have, this large selection ratio makes it challenging to narrow down to the best candidates. One fix is targeting efforts toward self-selection. He recalled an advertisement of the past that asked "Can you draw this?" to identify people with artistic talents. He asked if something could be developed and publicized to help people recognize if they have pattern recognition skills and whether they would be the right fit as a forensic examiner. Such an exercise could help fill applications with more of the right people that would not have to be filtered out.

Imada also commented on the notion of bias. It is not a weakness but rather a human condition, he said. In trying to accommodate for that human condition, labs need to recognize the existence of bias and cope with it organizationally and through the workplace culture, he stated. He noted that procedural strategies and events such as rotations, inspections, schedules, and work demands can either increase or mitigate the impact of bias depending on how they are managed.

[22] A funnel approach is a hiring strategy of multiple assessments to narrow a large pool of applicants to target candidates. See, for example, https://premierhrsolutions.wordpress.com/tag/funnel-approach/ [October, 2016].

He offered an impractical solution to hire people who lack the emotional need for satisfaction—what if examiners did not need to know the results of their work, that all they were doing was processing it. But, he said, "The fact is we all have feelings, we all have biases, and trying to eliminate [bias] is not as important as trying to recognize it and trying to minimize its effect on the kinds of work that we do."

Imada recognized that developing selection tests and improving the selection process of pattern evidence examiners is too large a problem for any one group of people or lab to do. The kind of funding and support required could come from an organization like the National Institute of Standards and Technology. He expressed that uncovering the critical knowledge, skills, aptitude, and other characteristics (KSAOs) and core competencies of this job and validation of an appropriate selection test would be extremely useful.

Becker reminded the audience that I-O psychology has a history of methodology to identify critical KSAOs and that information already exists useful for advancing selection in forensic science. She pointed to the Organizational Network website from the Department of Labor[23] and decades of research from the Dictionary of Occupational Titles. These resources have pages of detail on the tasks, knowledge, skills, abilities and other characteristics, including the motivations.

Ryan encouraged those in the forensic science field to capitalize on existing initiatives and methods. She was optimistic that it would be possible to validate some of the research presented at the workshop for selection purposes. Predictors of performance that already exist, like the form blindness test, could stand up to the scrutiny of the courts and be defensible as job relevant and consistent with business necessity if a formal validation study were conducted. Since many forensic labs are very small, Ryan agreed that a validation study could be conducted through a consortium or coalition of labs. She pointed out that in the private sector, selection tools are thought of as belonging to companies as part of their competitive advantage. But there is no real competition between labs in forensic science; people with the right skills and motivations in the field writ large are needed, and the nature of the profession is amenable to selection tools that can be used across labs.

Becker pointed out the synergy and sense of collaboration that developed at this workshop. Ryan recognized both the professional challenge and personal challenge to making sure the conversations continue. She said the workshop was an example of connecting different disciplines. Going forward, there are major conferences and other events where people with different perspectives and viewpoints might be invited. The personal challenge, suggested Ryan, is to create a network; add the new people that were met at the workshop and find ways to build on the information that could be useful.

In closing, Melissa Taylor (NIST) noted she has worked in the forensic science community for almost 15 years, dealing with the types of questions that came up in this workshop. (See Box 4-1 for a summary of the key questions and observations that were raised during the workshop.) She said she was pleased to observe the level of enthusiasm from both communities at the workshop in finding answers to tough and important questions. She thanked everyone for participating in not only the workshop, but also in the networking opportunities to

[23]See http://www.onetonline.org/link/summary/19-4092.00 [July, 2016] for a summary report on the job of forensic science technician with the tasks to collect, identify, classify, and analyze physical evidence related to criminal investigations. Other related job titles include: crime laboratory analyst, crime scene analyst, crime scene technician, crime scene investigator, evidence technician, forensic science examiner, forensic scientist, forensic specialist, latent fingerprint examiner, and latent print examiner.

create a community of interest that can grow in the future. She expressed her hope that this workshop would be a beginning of a new conversation on personnel selection in forensic science.

BOX 4-1
Key Questions and Observations Raised at the Workshop

What challenges do forensic laboratories face in hiring pattern evidence examiners?

- Hiring of examiners is often managed by a human resource unit separate from the labs and can be subject to uneven budget cycles and long waits.
- Evidence submitted to laboratories is rarely in perfect condition. Analysis of pattern evidence requires defining areas suitable for comparison and comparing samples to known source(s). This requires experts with specific visual and cognitive skills.
- There are long on-the-job training periods (about 2 years) before performance can be assessed.

What has been considered to address those challenges?

- Lab managers have observed on-the-job training to be more successful if new hires enter with a foundation of skills and cognitive abilities, notably visual acuity.
- Forensic laboratories are looking for mechanisms to detect necessary skills and cognitive abilities among applicants.
- Research has shown that individual differences exist on tasks of visual attention.

What future possibilities exist?

- There is a well-defined process for developing fair and appropriate selection tests in the field of I-O psychology that has been applied to many contexts for multiple professions and can be useful to the field of forensic science.
- Benefits of validated selection practices and assessments can include improved job performance, reduced turnover, reduced training costs, enhanced legal defensibility, etc.
- Some tests, like the form blindness test, show promise in detecting potential performance on visual tasks.
- Research on training is making strides toward understanding how individuals develop visual attention expertise.
- A formal validation study to develop selection test for forensic laboratories could capitalize on existing knowledge, methods, and tests.
- Given the size and resources of laboratories and the nature of the profession in the pattern evidence domain, a validation study could reasonably be conducted through a consortium or coalition of labs.

Appendix

Workshop Agenda and List of Participants

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

DIVISION OF BEHAVIORAL AND SOCIAL SCIENCES AND EDUCATION
Board on Human-Systems Integration

500 Fifth Street, NW
Washington, DC 20001
Phone: 202 334 3868
Fax: 202 334 2210
Email: bohsi@nas.edu
www.nationalacademies.org

A WORKSHOP
Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners

July 14-15, 2016

National Academies of Sciences, Engineering and Medicine
Room #100 Keck Center
500 Fifth Street N.W.
Washington, DC 20001

Note: This workshop is open to the public and will be recorded and webcast.

Workshop Objectives

- *To bring together industrial-organizational psychologists, experts on personnel selection and testing, forensic scientists, as well as other researchers whose work has a nexus with workforce needs in the forensic science field with a focus on pattern evidence.*

- *To develop a better understanding of the current status of selection and training of forensic scientists who specialize in pattern evidence, tools used in industrial and organizational psychology to understand elements of a task, and ways aptitude and performance can be measured.*

- *To discuss how these approaches could address challenges in the pattern evidence domain of the forensic sciences.*

Day 1

8:30 a.m.	*Registration Opens*

9:00 am **Welcome and Overview of Workshop**
Fred Oswald, Rice University, Committee Chair
Poornima Madhavan, BOHSI Director
Melissa Taylor, National Institute of Standards and Technology

9:20-9:30am **Status of Educational Preparation for Pattern Evidence Examiners**
Jay Siegel**,** Retired, Committee Member

9:30-10:00am **On-the-Job Experiences: The Task of Pattern Recognition**
Jessica LeCroy, Defense Forensic Science Center
Melissa Gische, Federal Bureau of Investigation

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

10:00-10:30am	**On-the-Job Experiences: Recruitment and Training** Maria Weir Ruggiero, Los Angeles County Sheriff's Department John M. Collins, Jr., The Forensic Foundations Group
10:30 am	*Break*
10:40-11:10am	**What is Industrial and Organizational Psychology? How might it be useful to forensic science?** S. Morton McPhail, Retired, President of the Society for Industrial and Organizational Psychology *Invited Discussant:* Winfred Arthur, Jr., Texas A&M University, Committee Member
11:10-11:45am	**Discussion**
11:45-12:45am	*Lunch*
12:45-1:15pm	**Test Development and Validation** Dan Putka, Human Resources Research Organization Nancy Tippins, CEB, Committee Member
1:15-2:15pm	**Discussion**
2:15pm	*Break*
2:15-2:45pm	**Attraction/Recruitment** Ann-Marie Ryan, Michigan State University, Committee Member Scott Highhouse, Bowling Green State University
2:45-3:45pm	**Discussion**
3:45-4:15pm	**Case Example** – Test Development, Considerations and Challenges Liberty Munson, Microsoft Learning Experiences
4:15-4:30pm	**Closing and Preview of Next Day** Fred Oswald, Rice University, Committee Chair
4:30 – 6:00 p.m.	*Day 1 Reception (Keck 100 and lobby)*

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Day 2

8:30 a.m.	*Registration Opens*
9:00 am	**Welcome, Review of Day 1, and Overview of Day 2** Fred Oswald, Rice University, Committee Chair Poornima Madhavan, BOHSI Director
9:10-10:30am	**Identifying Constructs** *What is known about pattern recognition? How is the state of the research advancing?* Zach Hambrick, Michigan State University, Committee Member Lisa Scott, University of Florida Bethany Jurs, Transylvania University Mara Merlino, Kentucky State University Mark Becker, Michigan State University
10:30am	*Break*
10:40-11:35am	**Pattern Evidence in the Courtroom** Randall Murch, Virginia Polytechnic Institute and State University, Committee Member Rockne Harmon, Retired Prosecutor Marvin Schechter, Defense Attorney Dan Murrie, University of Virginia *Invited Discussant:* Mara Merlino, Kentucky State University
11: 35am-12:15pm	**Discussion**
12:15-12:45pm	**Major Themes from Workshop** Wendy Becker, Shippensburg University Andrew Imada, A.S. Imada & Associates, Committee Member Ann-Marie Ryan, Michigan State University, Committee Member
12:45pm	**Closing Remarks** Fred Oswald, Rice University, Committee Chair

1:00 p.m. ADJOURN

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Division of Behavioral and Social Sciences and Education
Board on Human-Systems Integration

A Workshop
Personnel Selection in Forensic Science: Using Measurement to Hire Pattern Evidence Examiners

July 14-15, 2016

COMMITTEE MEMBERS

Frederick L. Oswald, Chair
Rice University

Winfred Arthur, Jr.
Texas A&M University

D. Zachary Hambrick
Michigan State University

Andrew S. Imada
A.S. Imada & Associates

Randall S. Murch
Virginia Polytechnic Institute and State University

Ann Marie Ryan
Michigan State University

Jay A. Siegel
Independent Consultant

Nancy T. Tippins
CEB

PRESENTERS

Mark W. Becker
Michigan State University

Wendy S. Becker
Shippensburg University

John M. Collins, Jr.
The Forensic Foundations Group

Melissa R. Gische
Federal Bureau of Investigation

Rockne P. Harmon
Independent Consultant

Scott Highhouse
Bowling Green State University

Bethany Jurs
Transylvania University

Jessica LeCroy
Defense Forensic Science Center

S. Mort McPhail
Society for Industrial and Organizational Psychology

Mara Merlino
Kentucky State University

Liberty Munson
Microsoft Learning Experiences

Daniel Murrie
University of Virginia

Dan Putka
HumRRO

Maria C. Ruggiero
Los Angeles County Sheriff's Department
Crime Laboratory

Lisa Scott
University of Florida

Marvin E. Schechter
Defense Attorney

Melissa Taylor
National Institute of Standards and
Technology

ADDITIONAL PARTICIPANTS

Susan Ballou
National Institute of Standards and
Technology

Heidi Eldridge
RTI International

Wesley Grose
Los Angeles County Sheriff's Department
Crime Laboratory

NATIONAL ACADEMIES STAFF

Poornima Madhavan
BOHSI Director

Julie Anne Schuck
Program Officer

Dixie Gordon
Information Officer

Annalee Gonzales
Senior Program Assistant