

PHOTO ARRAYS IN EYEWITNESS IDENTIFICATION PROCEDURES

By: Karen L. Amendola, PhD

POLICE FOUNDATION

Presented to:



NATIONAL ACADEMY OF SCIENCES

Committee on Scientific Approaches to Understanding and Maximizing the Validity and Reliability of Eyewitness Identification in Law Enforcement and the Courts



APRIL 24, 2014

MULTI-PART STUDY

- Emerged as part of the American Judicature Society's "Eyewitness Identification Field Studies" (see *Wells, Steblay, Dysart, 2011*)
 - The "partners" wanted to examine case outcomes and validate the lineup procedures and outcomes using a proxy for ground truth ("evidentiary strength")
 - Began after the AJS study was already underway in Tucson and Charlotte (but not in San Diego or Austin)
- Funded by multiple foundations (Jeht, Open Society Foundations, and Laura and John Arnold Foundation)

STUDY COMPONENTS:

FOLLOW UP TO THE AJS FIELD STUDIES

1. Development and validation of a rating instrument to assess evidentiary strength using key stakeholder groups;
2. Examination of the impact of presentation methods (sequential v. simultaneous) and pick types (no pick, filler pick, suspect pick) on case dispositions and ratings of evidentiary strength from the Wells, Steblay & Dysart (2011) study; and
3. Experimental study (Austin only):
 - A. Two teams of case evaluators (police, prosecutor, defense, and judge) assessed evidentiary strength with or without photo array information and outcomes
 - B. Examined impact of photo array outcomes on the interpretation of other case evidence (is there a biasing effect?)
 - C. Assessed differences in stakeholder groups' perspectives on evidentiary strength

INSTRUMENT DEVELOPMENT AND VALIDATION

- Brought together stakeholder groups in two separate two-day sessions
- Generated categories of evidence and evidence types
- Developed (individually and collectively) examples of weak, moderate and strong evidence within each type of evidence
- Process used was focused on developing content-oriented validity
- A survey of others not involved in original groups helped to further establish the rating scale “anchors” that were most psychometrically sound

INSTRUMENT VALIDATION (CONT'D)

- Stakeholder groups then evaluated 3 – 5 actual criminal case files using the instrument to pilot test its usefulness in creating greater objectivity within the subjective rating scale
- As part of the experimental study in Austin, we generated external validation evidence (concurrent) for the instrument

VALIDATION EVIDENCE

- The evidentiary strength ratings discriminated strongly between not prosecuted and adjudicated guilty cases

Table 7

Mean Overall Evidentiary Strength Ratings (1 – 5 scale) by Disposition within Pick Types

Photo array decision by victim/witness	Adjudicated Guilty	Not Prosecuted	Finding	<i>P</i>
No pick made	4.38	2.03	t(60)=11.25	≤ .001
Filler pick made	4.32	2.34	t(25)=6.331	≤ .001
Suspect pick made	4.33	3.14	t(45)=5.407	≤ .001

- Demonstrates initial criterion-oriented validity
- Instrument reliability and validity will be the topic of an upcoming article/monograph

STUDY ONE:

IMPACT OF PRESENTATION METHODS ON CASE OUTCOMES ACROSS STUDY SITES*

Presentation methods do not impact case dispositions
(n = 236)

Table 2

Frequencies of Case Dispositions by Lineup Presentation Methods Across Three Sites

Disposition	Sequential n (%)	Simultaneous n (%)	Total n (%)
Not Prosecuted	59 (59%)	88 (64%)	147 (62.3%)
Guilty (plea or judgment)	41 (41%)	48 (36%)	89 (37.7%)

$\chi^2 = .799$ (1 df), $p = \text{n.s.}$

*Excludes Charlotte due to early changes in protocol

SEQUENTIAL SUPERIORITY EFFECT?

Posterior probability of guilt higher with the simultaneous method.

Table 4

Differences in Case Disposition Proportions within Pick Type by Presentation Method Across Sites

Case Dispositions	Sequential n (%)	Simultaneous n (%)	Total n (%)	Chi square
No pick made				
Not Prosecuted	39 (72.2)	50 (78.1)	89 (75.4)	$X^2 = .551(1)$, n.s.
Guilty	15 (27.8)	14 (21.9)	29 (24.6)	
Suspect was picked				
Not Prosecuted	11 (34.4)	11 (29.7)	22 (31.9)	$X^2 = .170(1)$, n.s.
Guilty	21 (65.6)	26 (70.3)	47 (68.1)	
Filler was picked				
Not Prosecuted	9 (64.3)	27 (77.1)	36 (73.5)	$X^2 = .848(1)$, n.s.
Guilty	5 (35.7)	8 (22.9)	13 (26.5)	

SEQUENTIAL SUPERIORITY EFFECT?

The posterior odds of guilt when suspects are picked under the simultaneous method are 2.36 (26/11) versus 1.91 with the sequential method (21/11). As such, the observation that there is sequential superiority effect does not seem to be supported by these data, given that such an effect rests solely on a higher posterior probability of guilt (see Steblay et al., 2011).

STUDY ONE:

RELATIONSHIP BETWEEN PICK TYPES AND CASE DISPOSITIONS ACROSS STUDY SITES*

Suspect picks are associated with greater proportions of guilty findings. This is not a causal relationship.

Table 3

Frequencies of Case Dispositions by Pick Types Across Three Sites

Case Disposition	No Pick n (%)	Suspect n (%)	Filler n (%)	Total
Not Prosecuted	89 (75.4%)	22 (31.9%)	36 (73.5%)	147 (62.3%)
Guilty (plea or judgment)	29 (24.6%)	47 (68.1%)	13 (26.5%)	89 (38.1%)

$\chi^2 = 38.429$ (2 df) $p \leq .001$, Cramer's V = .404 $p \leq .001$

*Excludes Charlotte due to early changes in protocol

STUDY TWO: EXPERIMENT IN AUSTIN, TEXAS

- Raters randomly assigned to either receive or not receive information about the photo array and outcome (pick type)
 - Evaluations: Does knowledge of a photo array or its outcome influence the interpretation of other case evidence?
 - Does knowledge of a suspect pick inflate the overall evidentiary case strength rating?
 - Do photo arrays contribute to case dispositions?
 - Individual ratings by one police investigator, prosecutor, defense attorney, and judge
 - Discussion and final individual and group ratings
- Research questions:
 - Does knowledge of a photo array or its outcome influence the interpretation of other case evidence?
 - Does knowledge of a suspect pick inflate the overall evidentiary case strength rating?

RESULTS OF STUDY TWO: FINDING #1 (NON-ADJUDICATED CASES)

Inclusion of a photo array does **not** lead to significantly stronger case ratings.

Table 16

Overall Evidentiary Strength Ratings (1 – 5 scale) for Cases Not Prosecuted by Knowledge of Photo Array within Pick Types

Photo array decision by victim/witness	Knowledge of photo array across all rater types	No knowledge of photo array across all rater types	Finding and significance
No pick made	2.03	2.04	n.s
Filler pick made	2.34	2.56	n.s.
Suspect pick made	3.14	2.90	n.s.

RESULTS OF STUDY TWO: FINDING #2 (CASES WITH GUILTY FINDINGS)

Inclusion of a photo array **does** lead to significantly stronger case ratings but only for those cases in which suspects were picked and that were already rated exceptionally strong.

Table 17

Overall Evidentiary Strength Ratings (1 – 5 scale) by Knowledge of Photo Array for Cases Adjudicated Guilty within Pick Types

Photo array decision by victim/witness	Knowledge of photo array and outcome across all rater types	No knowledge of photo array across all rater types	Finding and significance
No pick made	4.38	4.51	n.s
Filler pick made	4.32	4.35	n.s.
Suspect pick made	4.33	3.99	t(64) = 2.275, p ≤ .05

RESULTS OF STUDY TWO: FINDING #3

Knowledge of photo arrays and their outcomes did not bias the interpretation of other case evidence by these criminal justice experts not even for the category labeled “Identification Information” suggesting that the other aspects of identifying information may play a larger role than the IDs themselves.

CONCLUSIONS

While IDs seemed to be diagnostic of guilt, they did not add anything to the overall interpretation of the case evidence as demonstrated in our experiment. This may be indicative of changes already occurring in the field and/or unique in that Austin has a policy generally restricting prosecution of ID cases without sufficient corroboration.