Law and Human Behavior, Vol. 25, No. 5, October 2001 (© 2001)

Eyewitness Identification in Actual Criminal Cases: An Archival Analysis[1]

Bruce W. Behrman[2,3] and Sherrie L. Davey[2]

This study analyzed 271 actual police cases in order to address several prevalent issues in the eyewitness literature. Suspect identification (SI) rates were obtained for 289 photographic lineups, 258 field showups, 58 live lineups, and 66 lineup identifications preceded by earlier identifications. SI rates were assessed for 3 levels of extrinsic evidence: no extrinsic evidence, evidence of minimal probative value, and evidence of substantial probative value. The SI rates for the photographic lineups were assessed as a function of delay, same vs. cross-race conditions, witness type, and weapon presence. SI rates declined significantly over time; SI rates were significantly greater for the same-race condition. SI rates were much greater for field showups than photographic lineups, 76% vs. 48%. The SI rates for the field showups did not vary as a function of eyewitness conditions. The relation between confidence and suspect/foil identifications for the live lineups was significant and moderately high. The utility of archival identification studies for eyewitness testimony research is discussed.

The research base in the field eyewitness testimony has been growing steadily for years, but there have been few studies that delve into real police cases. Because research in eyewitness memory has scarcely been tested against actual cases, it has been criticized for lacking external validity (Tollestrup, Turtle, & Yuille, 1994; Yuille, 1993). The scarcity of archival studies has led to the recognition of a need for a diversity of approaches to the problems of eyewitness memory, and in particular a need for studies that investigate real crimes. Archival work provides a complexity of variables not possible to simulate in the laboratory, but of course laboratory experiments afford specific controls that cannot be duplicated in real life. Thus, archival work should not be considered a standard against which experimental studies should be evaluated,

[1] Portions of this research were presented at the meetings of the American Psychological Society, June 1999, Denver, CO; and June 2000, Miami, FL.

[2] Department of Psychology, California State University, Sacramento, California.

[3] To whom correspondence should be addressed at Department of Psychology, California State University, 6000 J Street, Sacramento, California 95819.

0147-7307/01/1000-0475$19.50/1 © 2001 American Psychology-Law Society/Division 41 of the American Psychology Association

but rather as a valuable additional approach to the study of eyewitness issues. Several archival studies have been published that assess descriptions provided by witnesses in criminal cases: Kuehn (1974), Sporer (1992), and Tollestrup et al. (1994). Gonzalez, Ellsworth, and Pembroke (1994) assessed identifications made by witnesses to actual crimes as an adjunct to an experimental study on the suggestiveness of showups, but to date only one large-scale investigation has examined real-life identifications made under a variety of witnessing conditions (Tollestrup et al.). The present study investigates identification data from actual police cases in order to build upon the few archival studies already done.

A review of the literature indicates that several of the eyewitness issues examined here have been studied in great detail by means of experimental methods: the effect of delay, cross versus own-race effects, weapon focus effects, and the relation between confidence and accuracy. A large number of laboratory studies indicate that delay is an important factor in determining identification accuracy, and that the number of correct identifications declines as the interval between the crime and the identification procedure increases. (Deffenbacher, Carr, & Leu, 1981; Egan, Pittner, & Goldstein, 1977; Malpass & Devine, 1981a, 1981b; Shepherd & Ellis, 1973). The cross-race effect has been replicated in a large number of studies. These studies indicate that recognition memory is better for culprits of one's own race than for those of a different race (Anthony, Cooper, & Mullen, 1992; Meissner & Brigham, in press). One of the most widely studied eyewitness phenomena is that of weapon focus. Several studies have found that the presence of a weapon negatively affects a witness's ability to recognize the perpetrator of a crime. (Kramer, Buckhout, & Eugenio, 1990; Loftus, Loftus, & Messo, 1987). Summarizing this research, Steblay (1992) conducted a meta-analysis of various experimental studies of the weapon focus effect and found a difference between weapon-present and weapon-absent cases, with the weapon-present cases resulting in fewer accurate identifications than the weapon-absent cases. The literature in general indicates that the correlation between confidence and accuracy is low (Bothwell, Deffenbacher, & Brigham, 1987; Cutler & Penrod, 1989). However, more recent studies have found that substantial correlations may occur under certain conditions. For instance, Sporer, Penrod, Read, and Cutler (1995) have shown that moderate correlations may be obtained when only choosers (those witnesses who choose either the suspect or a foil) are included in the analysis. Some recent work by Lindsay, Read, and Sharma (1998) indicated that accuracy–confidence correlations were substantial when they were based on data collapsed across a variety of witnessing conditions. To date only the two issues of retention interval and weapon presence have been assessed by archival work and that too in only one study: Tollestrup et al. (1994). Tollestrup et al. examined both the effect of delay and weapon focus on the rate of suspect identification and obtained data consistent with the general thrust of the laboratory findings. In the present study we extended the Tollestrup et al. study and assessed suspect identification rates as a function of delay, and weapon focus; in addition, we examined same-race bias and the relation between suspect/foil identifications and confidence.

Tollestrup et al. (1994) have argued that the results of most laboratory studies pertain primarily to unaffected witnesses to a crime. These researchers argue

that victims of crimes and some witnesses to a crime have different personal and emotional levels of involvement in the crime than subjects in a laboratory experiment. Consequently, although laboratory results may apply to unaffected witnesses to a crime, it may be that crime victims are more accurate in their descriptions and also more likely to identify the suspect. Tollestrup et al. found that victims of robbery identified the suspect more often (46.5%) than did witnesses of robbery (33.3%). In a 1982 experimental study, Hosch and Cooper researched the relative accuracy of witnesses versus victims. However, the findings of their laboratory work indicated that being the victim of a crime does not significantly improve or decrease the ability to accurately identify a perpetrator. In light of the Tollestrup et al. arguments that unaffected witnesses may process information about the crime differently than victims who are more directly involved, we examined identification outcomes for both victims and witnesses to the crimes.

An important variable affecting eyewitness behavior focuses on the type of procedure used to identify suspects. Two basic types of identification procedures can be found in the literature, lineups, and field showups. A showup refers to the observation of a single suspect by a witness in the field, typically at the crime scene, whereas a lineup refers to the presentation of the suspect and several foils, either live or via photographs. Most of the controversy in this area of research has focused upon the suggestiveness of field showups. Some researchers have found that single suspect identification procedures (showups) result in more false identifications than lineups (Wagenaar & Veefkind, 1992; Yarmey, Yarmey, & Yarmey, 1996). Wagenaar suggests that witnesses in real life may trust the police to have arrested the correct person, resulting in an increased false alarm rate. The author warns against one-person identification procedures. A third study echoes the concerns of Wagenaar and Veefkind regarding showups (Malpass and Devine, 1983). Although Malpass and Devine do warn against biasing practices in lineups and photospreads, the authors believe that if a lineup is constructed fairly, with several foils that are similar looking to the suspect, it may be the preferred method of identification. The authors suggest that these procedures "... distribute the probability of identification of an innocent suspect across the lineup foils, reducing the risk of an identification error" (p. 83). In contrast to the previously discussed findings, Gonzalez et al. (1993) show that witnesses at a lineup are less likely to say "not there" than are witnesses at a showup. The study found that police suggestion is no greater for showups than for lineups. The study by Gonzalez and his colleagues is an important one in that currently showups are the more typical and preferred method of identification used by police departments (Gonzalez et al.). The current study distinguished between several types of identification procedures, and in particular assessed the differences between field showups and photographic lineups. In addition, we investigated the SI rates for single-photo presentations and live (physical) lineups.

Finally, we compared identification rates of suspects at lineups conducted after a previous identification procedure, either another lineup or a showup, with lineups administered with no prior identification. Courts have acknowledged that prior identifications may affect later identifications (*Simmons v. United States,* 390 U.S. 386 n. 6, 1968). Generally, if an earlier identification is considered to be biasing, the later identification will be considered tainted unless the witness states that he can

distinguish between his memory of the perpetrator and that of the suspect in the earlier identification (*People v. Orozco*, 1981). One commentator has noted that it is difficult to determine whether the viewer is identifying the perpetrator or the person observed in the earlier identification procedure (Sobel, 1971). Sobel is referring to the influence of photographs on later identifications, but the same reasoning applies to the earlier observation of a live person. Sobel has suggested that "there will be few instances where the witness will fail to identify the person most recently viewed . . ." (p. 296). However, no real-life evidence has been gathered that indicates that earlier identifications do indeed affect later ones.

A major problem with archival work is that a certain percentage of real world lineups do not include the perpetrator of the crime. Tollestrup et al. (1994) generated various evidentiary levels and asserted that such a procedure provided a solution to the problem of perpetrator-absent lineups. Like Tollestrup et al. we have attempted to alleviate the problem through the use of various categories of extrinsic incriminating evidence.[4] In this study we used three categories of evidence. In a large proportion of cases no evidence was recorded. A number of cases contained evidence that was incriminating but not particularly strong (e.g., the suspect has been implicated in prior cases with similar methods of operation). Finally, a number of cases included evidence that strongly implicated the suspect (e.g., a confession). Of course, one cannot be sure that the culprit is in the lineup or showup, even when the case includes highly incriminating evidence. Consequently, suspect identification rates should not be interpreted as measures of accuracy. Inferences regarding the convergence of archival and laboratory work should be drawn with caution. Archival identification research may be used to provide various measures of the choosing behavior of real witnesses and, viewed as such, may be used to explore areas of convergence with and divergence from laboratory results.

An additional problem may limit archival work. Tollestrup et al. (1994) noted that the police (the Royal Canadian Mounted Police in this case) did not distinguish between misidentifications and rejections of lineups. Either response was noted as a failure to identify anyone. We were able to gather only a small number of misidentifications for our photographic lineups, probably for the same reason. The Sacramento police officers generally note in their reports either that a witness identified the suspect, that the witness rejected the lineup (or showup) or was simply unable to make an identification. In several cases involving photographic lineups, detectives did note that a witness had chosen a foil and which one he had chosen, but these cases were comparatively rare. We were able to tabulate and record all false identifications for live lineups. Physical lineups are more formal procedures. Here the police recorded each pick made whether it was a suspect or one of the foils. In sum, most of the analyses reported in this paper were based on suspect identification rates; the analyses of live lineups included the calculation of misidentifications.

[4]Extrinsic incriminating evidence in this paper refers to incriminating evidence extrinsic to the identification of the person. In some cases suspects were identified not only by facial features but also by voice or body type, such identification evidence was not considered extrinsic. The identification of clothing was considered extrinsic to the identification, as were other more obvious forms of incriminating evidence, for instance, confessions or fingerprints.

METHOD

Descriptive Information

This archival analysis used files from the Sacramento City Police Department and a group of police reports from crimes committed in several counties in Northern California, including Sacramento County. A total of 271 cases were analyzed.[5] The total number of crimes was 349; the vast majority were armed robberies ($n = 261$). The remaining crime types included a variety of felonies: residential burglary, assault with a deadly weapon, car jacking, and attempted homicide. The crimes were committed between 1987 and 1998. Several types of eyewitness identification procedures were analyzed including 258 field showups, 289 photographic lineups, 58 live lineups, and 18 single photo showups. Most photographic lineups in our sample contained five photographs. All of the live lineups consisted of six people. We also analyzed 66 identifications that had been preceded by an earlier identification procedure. In addition to gathering demographic data on both the witnesses and the perpetrator, including gender, race, and birth date, each file was also assessed for date, time, and type of crime committed, weapon used, and whether the identifier was a victim or witness to the crime. Information on the identification type (including time and date), the number of suspect identifications, and confidence level (for the live lineups only) of the identification was gathered.

As already noted, the data for this study is based on 271 police cases. There were 374 perpetrators ($\mu = 1.38$; range: 1–7 per case), 282 victims ($\mu = 1.09$, range: 1–2), and 224 witnesses ($\mu = 0.83$; range: 0–6) . The average age of the male perpetrators was 27.24 years; the range was 18–68 years. Only three perpetrators were female; their ages were 23, 34, and 43 years. There were 161 female and 91 male victims.[6] The average age of the female and male victims was 33.15 and 35.36 years respectively; the range was 16–74 years for males and 15 to 70 years for females. There were 102 male witnesses and 84 female witnesses. The average age of male witnesses was 28.56; the range was 16–51 years. The average age of female witnesses was 34.10 years; the range was 15–67 years.

Evidence Categories and Booking Rates

We classified the cases in terms of extrinsic evidence levels in order to evaluate the validity of the lineup and showup identifications. These categories of evidence are listed in Table 1. We have divided the cases into two major groups: cases for which there is incriminating extrinsic evidence and cases for which no extrinsic evidence was recorded in the police reports. The cases containing extrinsic incriminating evidence were divided into two groups: those that included evidence of substantial

[5]The first author was consulted in 130 of these cases (These cases included cases from Sacramento County and several other counties in Northern California). The remaining 141 cases were researched at the Sacramento Police Department. The SI rates for these two groups of cases were not significantly different, either for field showups or photographic lineups. Seventy-three of the cases on which the first author was consulted proceeded to trial. The police reports themselves did not contain information regarding the further disposition of the cases. Conviction rates of the cases that proceeded to trial are not available.

[6]Gender data for 68 eyewitnesses was not recorded in the files. The age of 121 eyewitnesses was not recorded in the files.

Table 1. Extrinsic Incriminating Evidence Categories

Minimal probative value (MPV)	Substantial probative value (SPV)
Concurrent identification of clothing, weapon, or Vehicle[a]	Suspect on surveillance tape
Anonymously provided information; Information provided by cofelon[b]	Independent identification of weapon, clothing, or vehicle
Similarities of MO[c]	Stolen items, weapon(s), drugs found in suspects possession (on person; in residence or vehicle)
Previous familiarity with suspect[d]	Physical evidence: Ballistic evidence, fingerprints, electronically marked bills
	Confessions
	Vehicle at scene registered to suspect
	Identification of unique characteristics (tattoos)

Note: Reasons for classifying evidence as MPV are as follows: [a]Evidence identified at the same time and place as the suspect; subject to suggestion, [b]Subject to bias, [c]Relevance, [d]Subject to expectations.

probative value (SPV) and those that contained only evidence of minimal probative value (MPV). The first category (SPV) contains evidence that is of substantial probative value and is highly incriminating (e.g., confessions or fingerprint evidence). The second category (MPV) contains evidence that is minimally probative of guilt. We consider this a weaker form of evidence because it either suffers from relevance problems (e.g., similar MOs) or is subject to bias, suggestion, (e.g., identification of the clothing at a field showup while the clothing is still on the suspect), or the expectations of the witness (e.g., familiarity with the suspect). If a case contained both MPV and SPV types of evidence, we placed it in the SPV category. An MPV + SPV group was created by combining the MPV and SPV cases. The largest group of cases, the no extrinsic evidence (NEE) group, contained no evidence extrinsic to identifications of the suspect. We included in this group cases containing single or multiple identifications of the suspect, but excluded cases with extrinsic evidence as defined in Table 1.

We also noted the percentage of suspects booked for each class of evidence, as stated in the police reports. In California, arrest and booking for a felony is based on probable cause (California Penal Code, 1999), thus booking rates provided an independent measure of evidence level. We found that 92% of the suspects who were identified from photographic lineups and whose cases fit into the first category of evidence (SPV) were in fact booked into jail, whereas 71% of such suspects in the second category (MPV) were booked. The difference disappears for field-showups: 96% of the suspects whose cases fall into either class of evidence, SPV or MPV, were booked. Booking rates for all evidence categories for the photographic and field identification procedures are shown in Table 2.

ANALYSIS AND RESULTS

A total of 689 identification attempts from police cases ranging from homicide to nonviolent theft were analyzed. These included 258 field showups, 284 photographic lineups and 58 live lineups. A breakdown of suspect identification (SI) rates for various evidence and identification categories are presented in Tables 3 and 4.

Table 2. Booking Rates by Identification Procedure and Evidence Category

	NEE		MPV		SPV		MPV + SPV		Total	
ID type	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*
Photo	57	188	71	41	92	55	83	96	66	284
Field	77	184	96	23	96	51	96	74	83	258

Note: NEE denotes cases with no extrinsic evidence, MPV indicates cases with evidence that is of minimal probative value, and SPV indicates cases with evidence that is of substantial probative value.

Approximately 48% of the witnesses who observed a photographic lineup identified the suspect as the perpetrator. Delay of identification was a major factor affecting SI rates for the photographic lineups. Overall, 55% of the witnesses picked the suspect from a photo lineup if the identification attempt was within a 0–7-day delay period, whereas only 45% did so after a delay of greater than 7-days. The disparity between delay groups was particularly evident when the cases included hard (SPV) extrinsic evidence, 64 and 33% for the 0–7-day and 8+-day delay groups $\chi^2(1, N = 55) = 4.89$, $p < .05$, $\phi = .30$. When the two extrinsic evidence categories (MPV and SPV) were combined, the differences between the delay groups were significant, $\chi^2(1, N = 96), = 5.48$, $p < .05$, $\phi = .24$.

Two hundred and thirty-one witnesses participated in cross-race versus same-race identifications from the photographic lineups. The overall SI rate for the cross-race cases was 45%. In same-race identifications the SI rate was 60%. This difference was significant, $\chi^2(1, N = 231) = 4.50$, $p < .05$, $\phi = .14$. The extrinsic evidence categories (with the exception of the SPV category) exhibited similar trends, but the differences were not significant. The presence of a weapon did not reduce identification rates. Overall, the percentage of witnesses who selected the suspect in

Table 3. SI Rates and Evidence Categories for Photographic Lineups

	NEE		MPV		SPV		MPV + SPV		Total	
Condition	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*
Days delay										
0–7	49	55	75	8	64[a]	22	66[a]	30	55	85
8+	47	133	48	33	33	33	41	66	45	199
Race										
Cross-racial	42	117	64	14	46	28	52	42	45[b]	159
Intraracial	61	41	80	15	37	16	58	31	60	72
Victim vs. witness										
Victim	48	85	57	21	50	18	54	39	50	124
Witness	44	79	62	16	38	29	47	45	45	124
Weapon focus										
Weapon used	45	170	61	36	44	34	53	70	48	240
No weapon used	61	23	20	5	48	21	42	26	51	49

Note: NEE denotes cases with no extrinsic evidence, MPV indicates cases with evidence that is of minimal probative value, and SPV indicates cases with evidence that is of substantial probative value.

[a]For SPV, $\chi^2(1, N = 55) = 4.89$, $p < .05$; for SPV + MPV, $\chi^2(1, N = 96) = 5.47$, $p < .05$.

[b]For Total, $\chi^2(1, N = 231) = 4.50$, $p < .05$.

Table 4. SI Rates and Evidence Categories for Field Showups

	NEE		MPV		SPV		MPV + SPV		Total	
Condition	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*
Across conditions	78	184	71	51	78	23	73	74	76	258
Race										
Cross-racial	81	129	66	30	76	77	70	47	78	176
Intraracial	71	49	72	18	100	3	76	21	73	70
Victim v. witness										
Victim	73	86	76	29	67	9	74	38	73	124
Witness	80	89	60	20	82	11	68	31	75	120
Weapon focus										
Weapon used	74	155	74	38	81	16	76	54	75	20
No weapon used	96	27	67	15	71	7	68	22	82	49

Note: NEE denotes cases with no extrinsic evidence, MPV indicates cases with evidence that is of minimal probative value, and SPV indicates cases with evidence that is of substantial probative value.

crimes in which a weapon was involved was 48%, but in crimes without a weapon the percentage was 51%. None of the weapon/no-weapon differences for the various evidence categories was significant.

Victims performed slightly better than witnesses when the identification attempts were assessed as a whole, 50% versus 45%, but this difference was not statistically significant; nor were the differences for the various evidence categories.

We collected a total of 258 field showups. The average delay was less than a day. Ninety-three percent of the field showups were administered within a day after the crime. The SI rates are given in Table 4. Overall the SI rate was 76%. None of the differences in the race, weapon, or victim/witness categories were significant. The difference between the SI rates for the photographic lineups and the field showups (76% vs. 48%) was significant, $\chi^2(1, N = 542) = 45.19$, $p < .001$, $\phi = .29$. The difference between the SI rate for the photographic lineups administered 0–7 days after the crime and the field showups (76% vs. 55%) was significant, $\chi^2(1, N = 343) = 13.81$, $p < .001$, $\phi = .20$.

We collected data for a total of 58 live lineup identification attempts. All of the identification attempts for the live lineups reported here were administered by the Sacramento Police Department. Witnesses to live lineups administered by the Sacramento Police Department, fill out a standard identification form. The witness was asked to check one of three response options: (1) "I am sure that number ___ was the person who ... [committed the crime]," (2) "Although I am not positive, I think number ___ was the person who ...," and (3) "I did not recognize anyone in the lineup as being the person who ..." Consequently, for the live lineups we were able to count the SI rate, the false alarm rate, and the percentage of witness who marked "I did not recognize anyone ..." The SI rate was 50%, the false alarm rate was 24%; 26% of the witnesses were not able to make a choice. We computed χ^2 statistics and ϕ coefficients for the relationships between two dichotomous variables: (1) the choice of suspect or the choice of a foil versus (2) the statement that the witness was sure that the person was the culprit or the statement that the witness thought that the person was the culprit. A relatively high correlation was found between the degree of

confidence and whether a witness selected the suspect or a foil when all categories of evidence were combined, $\chi^2(1, N = 43) = 11.55$, $p < .001$, $\phi = .52$. Of the witnesses who expressed a moderate degree of confidence ("I think . . ."), the number of those who selected the suspect was 24, the number who picked a foil was 13. However, all but one of the witnesses who expressed a very high degree of confidence ("I am sure . . .") picked the suspect (19 picked the suspect, 1 picked a foil). We also computed the correlation between confidence and suspect/foil identification using the data from the combined extrinsic evidence categories (MPV + SPV). Here, the correlation was moderately high, $\chi^2(1, N = 34) = 8.09$, $p < 01$, $\phi = .49$. In this case, 15 of those who were "sure" of their pick chose the suspect, 1 chose a foil. Of those who expressed a moderate degree of confidence ("think") 19 chose the suspect, 10 chose a foil. Finally, we computed the correlation between confidence and suspect/foil picks using the data only from the highly incriminating evidence category (SPV). For this data, the correlation was again moderately high, $\chi^2(1, N = 27) = 6.07$, $p < .05$, $\phi = .47$. A very similar pattern of responses was evident: Twelve of the "sure" witnesses picked a suspect, only 1 chose a foil; 15 of the "think" witnesses chose the suspect, 8 chose a foil.

Our data contained 18 single-photograph identifications. These are analogous to field showups, yet we thought them to be sufficiently different to warrant a separate analysis. The SI rate for the single photos was .83. The average delay was less than a day. The SI rate is consistent with that of the field showups.

We computed the SI rates for 66 identifications that were preceded by earlier identification procedures. The later lineups were either photographic or live lineups. The earlier identifications were based on either field showups or photographic lineups. The average delay period for the earlier identifications was 11.43 days. The average delay for the later identification was 28.18 days. The SI rate for the later identifications was 62%. This rate was compared with an SI rate computed from a group of live ($n = 26$) and photographic ($n = 199$) lineups that had not been preceded by earlier identifications, 225 identification attempts in all, and that had occurred 8 or more days after the crime. The SI rate for this control group was 45%. The difference between the control group and group that had observed the suspect at an earlier identification procedure was significant, $\chi^2(1, N = 291) = 6.06$, $p < .05$, $\phi = .14$. We also assessed the number and type of responses for both the earlier and later identifications. Of the 33 witnesses who identified the suspect at the earlier procedure, 24 again identified him at the later one and 9 were not able to identify anyone at the later identification.[7] Of the 31 witnesses who were not able to make an identification at the first identification procedure, 17 were not able to make an identification later, but 14 did identify the suspect at the later identification. This means that 27% of the witnesses who made a positive earlier identification did not make a later identification, whereas 45% of the witnesses who did not identify the suspect at the earlier identification procedure did so at the later one. Finally, we computed the

[7]Some police reports did indicate that a witness had picked a foil from a photographic lineup. However, only a few reports indicated that a foil had been chosen. Consequently, we placed foil choices in a "no identification" category for the purpose of this analysis. Thus, responses in the "no identification" category were of three types: lineup rejections, choosing of foils, and responses indicating that the witness was not sufficiently certain to make a choice.

correlation between the SI rate and the length of delay between the first and second identifications.[8] The delay was divided into two intervals, delays from 0 to 30 days, and delays greater than 30 days. Of the 43 later responses that occurred 30 or fewer days after the first identification, 60% were positive identifications of the suspect; of the 23 later responses that followed the first identification by 31 or more days, 65% were positive identifications of the suspect. The correlation between the SI rate and delay was not significant, $\chi^2(1, N = 66) = .14$, $p > .05$, $\phi = 05$.

DISCUSSION

As expected, delay of identification was a major factor affecting SI rates for the photographic lineups. Our results indicate that the SI rate drops substantially after 7 days. This is true of all extrinsic evidence categories, MPV, SPV, and MPV + SPV. Thus, our findings converge with the laboratory evidence, which has, with few exceptions, indicated that the rate of identifying the culprit declines with time (Shapiro & Penrod, 1986). The decline in SI rate over time may invite the conclusion that recognition memory degraded over time, but it would be wrong to draw such a blanket conclusion from these data. A memory component is likely responsible for a portion of the decline in the SI rate. However, a host of variables influences witnesses' decisions in real cases. For example, witnesses to real crimes may simply become more cautious as the delay between the crime and the identification increases. Our results are also consistent with those of Tollestrup et al. (1994). We did not find a decline over the first few (0–2) days as did Tollestrup et al. We computed SI rates for each day for delays up to 30 days. The rates were similar, hovering around 60–70% (with the exception of the NEE category) for each of the days within the first week, and bottomed out after the first week at about 40%.

The tendency to choose the cross-racial suspect from the photographic lineups is lower than the tendency to pick the intrarace suspect in all evidence categories, except the SPV category. These results are in line with the effects found by many in the field (Malpass & Kravitz, 1969; Malpass, Laviqueur, & Weldon, 1973; Platz & Hosch, 1988; Shepherd, Deregowski, & Ellis, 1974). Overall, our findings are consistent with the meta-analytic work on cross-race effects indicative of a small-to-moderate decrease in the rate of identifying the culprit when the identification is cross-racial (Anthony et al., 1992; Meissner & Brigham, in press). A reversal occurs in the SPV group; this is puzzling and at odds with the literature in general. As already noted, the perpetrator is more likely to be present in the SPV group. It is possible that witnesses tend to pick a suspect of a different race when there is even a vague memory of the perpetrator available to them and when the suspect is in fact the perpetrator. It may be that a witness's decision criterion is somewhat less strict for individuals of a different race (Slone, Brigham, & Meissner, 2000), and that this tendency to choose other-race faces manifests itself when the perpetrator is present in the lineup. Of course the number of witnesses in the SPV group is small, and the difference between the same and cross-race SPV groups are not statistically significant. In general our data indicate

[8]We thank an anonymous reviewer for suggesting that we compute this correlation.

that witnesses do not pick suspects from a different race more often than those of their own race.

A host of studies have supported the phenomenon of weapon focus, including meta-analytic studies (Kramer et al., 1990; Loftus et al., 1987; Steblay, 1992); however, support for the weapon focus effect was not evident in the present study. Further, the suspect was chosen more often in the weapon-present group than in the weapon-absent group for both the MPV and MPV + SPV conditions. One possible explanation for our finding is that weapon-focus is simply not a real-life phenomenon. Perhaps the arousal level during a crime is simply too high in real life crimes, whether a weapon is present or not. The emotions of fear or anger, associated with weapon-present crimes, are not present to the same extent in a laboratory as in a real crime (Yuille, 1993). Still, this explanation does not explain the difference between the results of our archival study and the archival study of Tollestrup et al. (1994) that corroborated the laboratory findings. It should be pointed out, however, that when an analysis of covariance was performed on the Tollestrup et al. data the presence of a weapon had only a marginally significant effect on the suspect identification rate.

Our findings do not indicate that victims respond differently than witnesses in identifying suspects. Tollestrup et al. (1994) and Yuille (1993) have suggested the emotional arousal experienced by victims might result in different response levels for victims than for unaffected bystanders or even victims of nonviolent crimes, for example, fraud cases. Tollestrup et al. did find a significant association between witness types (robbery victims, robbery witnesses, and fraud victims) across evidence levels. Tollestrup et al. found that SI rates were related to witness type. Across evidence categories the SI rates were 46.5, 33.3, and 25.5 for robbery victims, robbery witnesses, and fraud victims respectively. We found the SI rate only slightly higher for victims than witnesses overall. The SI rate was higher for victims than for witnesses in both the SPV and the combined extrinsic evidence categories, MPV + SPV, although the percentages were reversed for the MPV group. None of these differences were significant. We agree with Tollestrup et al. that a victim should be defined as one who "directly interacts" (p. 146) with the perpetrator. We evaluated the victims and witnesses as designated in the police reports in terms of Tollstrup's "direct interaction" criterion and found that victims did indeed experience a more direct interaction with the perpetrator than the witnesses. However, in many of our cases a witness may have experienced a high degree of emotion even though they had not been directly involved (e.g., a person in line during bank robbery as compared with the teller who had been held up by the robber). Thus our sample of witnesses may have included witnesses who were quite emotionally affected by the crime even though they had not directly interacted with the perpetrator. The fact that our witness group may have contained witnesses who were highly emotional may account for the fact that our data do not clearly support the Tollstrup et al. results. In addition, it should be noted that Tollestrup et al. found a significant association between witness types and identification outcomes when comparing robbery victims, robbery witnesses, and fraud victims, but a subsequent post hoc multiple comparison test did not reveal significant differences between these groups.

The relationship between confidence and suspect/foil identification for the live lineups is a solid one; indications are that highly confident witnesses choose fewer distracters than moderately confident ones. We obtained significant correlations for the highly incriminating evidence category (SPV), the inclusive evidence category (MPV + SPV), and for the data computed across all categories of evidence. Only one highly confident witness in the sample of witnesses at live lineups picked a foil. Wells and Bradfield's (Wells & Bradfield, 1998) observation that detectives may inadvertently cue witness when they know who the suspect is may be correct in certain cases; such inadvertent cuing may have happened in some of our cases. It must be noted however that the Sacramento Police Department is aware of the effects of bias on the part of detectives. According to the Identification Administrator for the SPD (H. Ayers, personal communication, September 10, 1997) detectives are instructed to be careful concerning any statements or actions they might make during or after the identification, which might be considered suggestive. The SPD detectives are instructed in the contents of *The California Police Officers Legal Sourcebook* (California Department of Justice, 1993), including the chapter on lineups and showups. Section 8C of that chapter states: "... if you tell a witness that he has picked the 'right' or ('wrong') person, it may jeopardize the admissibility of later in-court identifications." (p. 229). Of course the detectives present at the lineup may still unconsciously cue witnesses, but at least their awareness of the problem may mitigate some of the suggestiveness inherent in the post-lineup situation.

The bulk of the experimental literature indicates that the confidence–accuracy relation is a weak one (Sporer et al., 1995). However, a recent trend in the experimental literature indicates that under some circumstances confidence may be a good predictor of accuracy. Lindsay et al. (1998) obtained confidence–accuracy correlations of .51 and .68 for two different videos when the witnessing conditions varied widely. Sporer et al. (1995) found a substantial confidence–accuracy correlation when the analysis was limited to witnesses who made positive identifications. Our results parallel those of Lindsay et al. (1998) and Sporer et al. (1995); however, it would be wrong to infer that our data indicate that confidence predicts accuracy. Our findings do indicate that highly confident witnesses are much more prone to choose the suspect in a criminal proceeding than are moderately confident ones. Further, it is more likely that the witnesses' choices are correct in the SPV cases than other cases. It also should be emphasized that our confident witnesses (except for one) did not make misidentifications, and in this sense were seldom inaccurate.

The data from the field showups is inconsistent with those obtained by Gonzalez et al. (1994), which indicated that field showups produced fewer suspect identifications than lineups. Two of the studies reported in the Gonzalez et al. article were laboratory studies; the SI rates, in both the perpetrator-present lineups and the perpetrator-absent lineups were greater for lineups than the showups. Study 3 of the Gonzalez et al. report was an archival study. The percentage differences were striking. Only 22% of the witnesses positively identified the suspect at the showups, whereas 75% did so when photographic lineups were used. It should be noted that the sample used to compute these lineups figures was small ($n = 24$). Our findings show that the suspect identification rate is highest for field showups. Our results indicate that field showups produced a high rate of suspect identification, a finding that

is consistent with the hypothesis that showups are suggestive. The SI rates were 76% for field showups and 48% for photographic lineups, $n = 258$ and 284 respectively. These differences carry throughout all of the evidence levels. Of course, some of the differences reported in our study may be confounded by the amount of delay. However, we found that the SI rate for lineups administered within the first week were significantly smaller than those for showups. The relatively high SI rate for the lineups reported by Gonzalez et al. may in part be due to the small sample, but it is the low SI rate (22%) for showups, which is troubling. The Gonzalez et al. sample of showups is large ($n = 172$). At this point we have no explanation for the differences between the two studies. At any rate, our results are consistent with the laboratory findings that showups are biasing (Yarmey et al., 1996). In addition, our study confirms the intuition of the United States Supreme Court that the risks of bias are greater in showups than lineups (*Stovall v. Denno*, 1967). Several other courts have expressed the opinion that showups are suggestive (*People v. Johnson*, 1989; *People v. Orozco*, 1981). In fact, this opinion is expressed in the National Police Foundation report, *Model Rules: Eyewitness Identification* (Project on Law Enforcement Policy and Rulemaking, 1974), and in the U.S Department of Justice document, *Eyewitness Evidence: A Guide for Law Enforcement* (Technical Working Group for Eyewitness Evidence, 1999), both of which indicate that showups are biasing, and suggest methods to mitigate the bias. It might be noted that the booking rate for the showups in our study for which there is no extrinsic evidence is 77%, whereas the rate for the extrinsic evidence categories is about 96%; nevertheless, the rate of identification for the no-evidence category is 78%, equal to or slightly higher than it is for those cases that include extrinsic incriminating evidence. Certainly, booking rates are not a failsafe method of assessing perpetrator-absent identifications, but it seems reasonable to argue that a showup category with a 23% nonbooking rate contains more perpetrator-absent showups than does a category with a 4% nonbooking rate. Consequently, the fact that the NEE category has a *higher* SI rate than the SPV or the MPV + SPV extrinsic evidence categories provides additional support for the view that showups are indeed biasing.

The field showup data was assessed in terms of both evidence categories and the typical eyewitness categories of race, weapon presence, and victim/witness differences. In general, the SI rates were high for the various groups. The cross-race groups produced lower SI rates than the own-race groups for all incriminating evidence categories, a trend consistent with the experimental data. The weapon-present groups in the MPV, SPV, and the MPV + SPV evidence categories produced higher SI rates than the weapon-absent groups, a trend not expected from the laboratory data. None of these categories produced significant differences. It may be that the images generated in field showups are too fresh and vivid to allow for the variables of race, weapon presence, or witness type to produce large, statistically significant effects.

The current study clearly indicates that witnesses at field showups react differently than those assessing a photographic lineup. We have found that witnesses demonstrate high SI rates at field confrontations. But perhaps of greater interest is the fact that none of the classic eyewitness factors, race, weapon presence, or witness type, produced significant or even marginally significant effects when the identifications were made at field showups. If most jurisdictions identify suspects primarily

through the use of showups, (Gonzalez et al., 1993) many of the factors found to be influential in the laboratory are irrelevant in a large proportion of real cases. Certainly later identifications might be based on photographic or live lineups, but these would likely be tainted by the earlier field identification. The finding that SI rates do not vary with many of the eyewitness factors that influence witness behavior at lineups does not mitigate the biasing effect of field showups. However, it does underline the point that the eyewitness literature may have to take into account the methods of identification to a far greater extent than in the past when assessing the effects of situational or witness variables on identification.

The data comparing the identification rates of suspects at lineups conducted after witnesses have observed the suspect at an earlier procedure with lineups conducted with no prior identification procedure, are consistent with the hypothesis that earlier identifications may contaminate later ones. Further, the data tracking the changes in "yes" and "no" responses (which included rejections, "I don't know" responses and false alarms) indicate that a process similar to the concept underlying unconscious transference is operating here (Ross, Ceci, Dunning, & Toglia, 1994). A large number of witnesses who say "no" or "I don't know" in the earlier identification say "yes" later. Apparently, these witnesses did distinguish, at least to some degree, the perpetrator and the suspect at the earlier lineup. Either the witness recalled that the person in the preliminary identification procedure did not fit their image of the perpetrator or their image of the perpetrator was simply too vague to say "yes." Evidently, whatever image the witnesses had of the suspect at an earlier time became more dominant at the later identification. The witness's image of the perpetrator could have been overwritten by the image of the suspect in the earlier identification procedure; in other cases the witness could have confused the sources of the images (Ross et al., 1994). If a witness rejected the earlier identification because they had two distinct images, that of the perpetrator and that of the identification picture, then at the later identification procedure he may have recalled the image in prior lineup as the one at the crime scene. Whether the process involved is one of overwriting or one of source exchange, the data indicate that the image observed at the earlier identification has become the one associated with the perpetrator. It would then be difficult for a witness to any longer distinguish his memory of the perpetrator from that obtained at the preliminary identification. In either case these data are consistent with the laboratory findings on bystander misidentification. These findings do argue against putting a suspect in a later lineup, when he has been observed at an earlier one. Chances of identifying the suspect are greatly increased and this is the case even when the later identification is administered some months after the crime. Any argument that a live lineup could be used for elimination purposes after the suspect has once been identified is highly questionable in light of this data.

There are several limitations to this study in need of mention. The cases cover a broad range of crimes. Our set of crimes was similar to those of the Tollestrup et al. (1994) study in that our cases included a wide variety of robberies; however, our set was slightly broader than Tollestrup et al.'s in that it contained additional types of crimes, for example, residential burglaries. The essential aspects of the crimes in our study were that the victims were aware that a crime was taking place, and that the observation times were relatively short. The victims in our cases were quite

likely attentive to the perpetrator, unlike many victims of fraud (Tollestrup et al.). Consequently, our results should not be applied to cases involving fraudulent activity in which the victims may not realize the culprit is a criminal until long after the event. There were certainly a wide range of viewing conditions and viewing times in our study; still, the crimes did not provide lengthy opportunities for rehearsal as did, say, the homicide assessed in Yuille and Cutshall's seminal field study (Yuille & Cutshall, 1986). In the current study, the crimes occurred relatively quickly, but future work might vary viewing conditions such as time to observe, in order to assess the interaction of viewing times with other situational factors.

It was difficult to obtain large samples for a number of analyses. We found it difficult to achieve large samples of cases in either the MPV or the SPV categories. Weapon-absent cases were particularly difficult to come by. Another problem was the difficulty in obtaining live lineups. These are relatively infrequently administered in California and, we gather, in other jurisdictions (Tollestrup et al., 1994). Thus, our sample of live lineups was restricted. This in turn limited our sample of confidence ratings, because a standard scale of confidence is administered only for live lineups by the Sacramento Police Department. Many expressions of confidence were recorded in the police reports for the photographic lineups, a very large category of identifications; however, as noted previously, the police reports contained only a few instances of recorded misidentifications, making it impossible to obtain a large sample of confidence–suspect/foil relations.

The current study clearly demonstrates the validity of comments by Tollestrup et-al. (1994) to the effect that forensic research is multifaceted and that factors affecting eyewitness behavior may vary from one context to another. One of the problems with the literature to date has been the assumption that eyewitness factors have the same effect and that witnesses respond similarly regardless of the forensic situation (Tollestrup et al., 1994). A diversity of methods is needed if we are to provide the legal profession with practical advice regarding eyewitness memory. It is from a combination of methods, controlled experiments, field studies, and archival studies that conclusions should be drawn. An examination of the similarities and differences between the identification outcomes generated by the different approaches may determine when one approach is applicable in contexts other than its own. The results of the study reported here converge with several eyewitness findings from the laboratory: the effect of delay and of cross-race identification, and the biasing effect of showups; however, our results diverge from the experimental literature with regard to weapon focus effects, and to some extent the relationship between confidence and suspect identification. Thus, the present study highlights the need for several approaches to eyewitness research.

ACKNOWLEDGMENTS

We thank the Sacramento Police Department (SPD) for providing access to its police reports on major crimes and to the Records Department of the SPD for assisting us in researching the files. We especially thank Capt. M. Busch and Harold Ayers of the SPD for their invaluable assistance with this research. We thank Carole

W. Barnes for providing the initial liaison with the SPD. We thank Susan Kollenberg and Shannon Corey for their time and effort in categorizing and collating the massive amount of data collected. We thank Nadia Dauod, Rita Mechill, and Mona Stilson for their extraordinary contributions to the collection of data for this research. Critical comments of Betsy Feldman on an earlier version of the article are gratefully acknowledged.

REFERENCES

Anthony, T., Cooper, C., & Mullen, B. (1992). Cross-racial facial identification: A social cognitive integration. *Personality and Social Psychology Bulletin*, *18,* 296–301.

Bothwell, R., Brigham, J., & Malpass, R. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin*, *15,* 19–25.

Bothwell, R., Deffenbacher, K., & Brigham, J. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72,* 691–695.

California Department of Justice. (1993). *California Peace Officer's legal sourcebook.* Sacramento: Author.

California Penal Code, West's Annotated California Codes: Penal Code S 836 West Group 1999.

Cutler, B., & Penrod, S. (1989). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology*, *74,* 650–652.

Deffenbacher, K., Carr, T., & Leu, J. (1981). Memory for words, pictures and faces: Retroactive inference. *Journal of Experimental Psychology, Human Learning and Memory*, *7,* 299–305.

Egan, D., Pittner, M., & Goldstein, A. (1977). Eyewitness identification: Photographs vs. live models. *Law and Human Behavior*, *1,* 199–206.

Gonzalez, R., Ellsworth, P. C., & Pembroke, M. (1993). Response biases in lineup and showups. *Journal of personality and Social Psychology*, *64,* 525–537.

Hosch, H. M., & Cooper, D. S. (1982). Victimization as a determinant of eyewitness accuracy. *Journal of Applied Cognitive Psychology*, *67,* 649–652.

Kramer, T., Buckhout, R., & Eugenio, P. (1990). Weapon focus, arousal and eyewitness memory. *Law and Human Behavior*, *14,* 167–184.

Kuehn, L. L. (1974). Looking down a gun barrel: Person perception and violent crime, *Perceptual and Motor Skills*, *39,* 1159–1164.

Lindsay, S. D., Read, J. D., & Sharma, S. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9,* 215–219.

Loftus, E., Loftus, G., & Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behavior*, *11,* 55–62.

Malpass, R., & Devine, P. (1981a). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, *66,* 482–489.

Malpass, R., & Devine, P. (1981b). Guided memory in eyewitness identification. *Journal of Applied Psychology*, *66,* 343–350.

Malpass, R. S., & Devine, P. G. (1983). Measuring the fairness of eyewitness identification lineups. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 81–102). New York: Wiley.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13,* 330–334.

Malpass, R., Laviqueur, H., & Weldon, D. (1973). Verbal and visual training in face recognition. *Perception and Psychophysics*, *14,* 285–292.

Meissner, C. A., & Brigham, J. C. (in press). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law.*

People v. Johnson, 210 Cal. App. 3d. 316 (1989).

People v. Orozco, 114 Cal. App. 3d. 435 (1981).

Platz, S., & Hosch, H. (1988). Cross-racial/ethnic eyewitness identification: A field study. *Journal of Applied Social Psychology*, *18,* 972–984.

Project on Law Enforcement Policy and Rulemaking. (1974). Model rules: Eyewitness identification. Washington, DC: Police Foundation.

Ross, D. F., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and mistaken identity: When a witness identifies a familiar but innocent person. *Journal of Applied Psychology*, *79,* 918–930.

Shapiro, P., & Penrod, S. (1986). Meta-Analysis of facial identification studies. *Psychological Bulletin*, *100,* 139–156.

Sheperd, J. W., Deregowski, J. B., & Ellis, H. D. (1974). A cross-cultural study of recognition memory for faces. *International Journal of Psychology*, *9,* 205–211.
Shepherd, J., & Ellis, H. (1973). The effect of attractiveness on recognition and memory for faces. *American Journal of Psychology*, *86,* 627–633.
Simmons v. United States, 390 U.S. 377 (1968).
Slone, A. E., Brigham, J. C., & Meissner, C. C. (2000). Social and cognitive factors affecting the own-race bias in whites. *Basic and Applied Social Psychology*, *22,* 71–84.
Sobel, L . (1971). Assailing the impermissible suggestion, *38 Brooklyn Law Review*, 263–271.
Sporer, S. L. (1992). *Person descriptions in an archival analysis of criminal cases.* Unpublished manuscript, University of Marburg, Germany.
Sporer, S. L., Penrod, D., Read, S., & Cutler, B. (1995). Choosing, confidence and accuracy: A meta-analysis of the confidence relation in eyewitness identification studies. *Psychological Bulletin*, *118,* 315–327.
Steblay, N. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, *16,* 413–424.
Stovall v. Denno, 388 U.S. 293 (1967).
Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement.* Washington DC: U.S. Department of Justice, Office of Justice.
Tollestrup, P., Turtle, J., & Yuille, J. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 144–160). New York, NY: Cambridge University Press.
Wagenaar, W. A., & Veefkind, N. (1992). *Comparison of one-person and many-person lineups: A warning against unsafe practices.* Unpublished manuscript, Leiden University, Leiden, The Netherlands.
Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect" Feedback to the witness distorts their reports of the witnessing experience. *Journal of Applied Psychology*, *83,* 360–376.
Yarmey, D. A., Yarmey, M. J. , & Yarmey, A. L. (1996). Accuracy identifications in showups and lineups. *Law and Human Behavior*, *20,* 459–477.
Yuille, J. C., (1993). We must study forensic eyewitnesses to know about them. *American Psychologist*, *48,* 572–573.
Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory for a crime. *Journal of Applied Psychology*, *71,* 291–301.