

*Psychology, Department of*  
*Faculty Publications, Department of*  
*Psychology*

---

University of Nebraska - Lincoln

Year 2006

---

Mugshot Exposure Effects: Retroactive  
Interference, Mugshot Commitment,  
Source Confusion, and Unconscious  
Transference

Kenneth A. Deffenbacher\*

Brian H. Bornstein†

Steven D. Penrod‡

\*University of Nebraska–Omaha, kdefenbacher@mail.unomaha.edu

†University of Nebraska-Lincoln, bbornstein2@unl.edu

‡John Jay College of Criminal Justice, New York, spenrod@jjay.cuny.edu

This paper is posted at DigitalCommons@University of Nebraska - Lincoln.

<http://digitalcommons.unl.edu/psychfacpub/175>

## Mugshot Exposure Effects: Retroactive Interference, Mugshot Commitment, Source Confusion, and Unconscious Transference

Kenneth A. Deffenbacher, Department of Psychology, University of Nebraska–Omaha  
(Corresponding author: e-mail: [kdeffenbacher@mail.unomaha.edu](mailto:kdeffenbacher@mail.unomaha.edu))

Brian H. Bornstein, Department of Psychology, University of Nebraska–Lincoln

Steven D. Penrod, John Jay College of Criminal Justice, New York

*Published online May 31, 2006*

**Abstract:** More than 25 years of research has accumulated concerning the possible biasing effects of mugshot exposure to eyewitnesses. Two separate metaanalyses were conducted on 32 independent tests of the hypothesis that prior mugshot exposure decreases witness accuracy at a subsequent lineup. Mugshot exposure both significantly decreased proportion correct and increased the false alarm rate, the effect being greater on false alarms. A mugshot commitment effect, arising from the identification of someone in a mugshot, was a substantial moderator of both these effects. Simple retroactive interference, where the target person is not included among mugshots and no one in a mugshot is present in the subsequent lineup, did not significantly impair target identification. A third metaanalysis was conducted on 19 independent tests of the hypothesis that failure of memory for facial source or context results in transference errors. The effect size was more than twice as large for “transference” studies involving mugshot exposure in proximate temporal context with the target than for “bystander” studies with no subsequent mugshot exposure.

**Keywords:** mugshots, photobiasing lineups, retroactive interference, source confusion, unconscious transference

The phenomenon of the photobiasing lineup has existed for as long as police have, in the furtherance of their criminal investigations, engaged in the practice of exposing eyewitnesses to mugshots. If a witness is exposed to mugshots subsequent to viewing a perpetrator and prior to an additional test of recognition memory, there is the possibility that exposure to the mugshot photos may bias the witness’s decision at that test, usually a photo or live lineup. Defense attorneys have labeled a photobiasing lineup as one in which a possibly innocent person has been arrested because his/her photo was initially included in a set of mugshots but not identified

by the witness; subsequently, however, he/she was identified by that witness at a lineup. Any bias thus engendered would be due to the previous exposure of the witness to the mugshot photo of the suspect identified at the subsequent lineup. In such cases, the witness may well have suffered a failure of memory for the circumstances of previous encounter of the face identified at the lineup. This sort of failure of memory for facial source or context is all the more problematic when viewing of the perpetrator has occurred under less than optimal viewing conditions (Brown, Deffenbacher, & Sturgill, 1977). Indeed, the United States Supreme Court had opined that this was the case, even before there was empirical support for their view (*Simmons v. United States*, 1968).

The first empirical examination of the possibility of bias imposed by mugshot exposure more generally as well as the more specific sort of bias that might be engendered by a photobased lineup was that of Brown *et al.* (1977). In all three of their experiments, these investigators noted that recognition memory for a once-seen face is much greater than is the capacity to recall the circumstances under which that previously unfamiliar face had been encountered. In Experiments 2 and 3, they employed what we will henceforth refer to as a *transference design* to test for specific biasing effects of mugshot presentation between exposure to live targets and a subsequent lineup.

A transference design occurs in any study in which the ultimate test of recognition memory includes a face of a person who shared proximate temporal context with the target person. Two sorts of proximate temporal context have been investigated with the use of a transference design, one where the "proximate person" was seen in mugshots shortly after encounter with the target person and one where the "proximate person" was encountered as a bystander to the target event. For instance, Read, Tollestrup, Hammersley, McFadzen, and Christensen (1990, Exp. 5) have shown that proximate temporal context for a target face (a sound technician viewed by everyone for 4–5 min in an introductory psychology class) and a "bystander" face (a male bystander who had been seen by half the students in one of several roles, distributing exam materials in other classes, for instance) can be as much as 1 week. That is, they have shown that exposure to target and bystander on the same day is equally productive of transference bias as is exposure to a bystander 1 week after exposure to the target. It could well be that exposure to mugshots as much as 1 week after exposure to the target would be productive of transference bias, too.

At any event, there is a transference effect if a previously seen nontarget face is chosen from a fair lineup, ordinarily a target-absent (TA) one, at a significantly greater rate than for previously unfamiliar foils. Consider a fair target-absent lineup of five persons. If the tendency for witnesses to pick the previously seen innocent person were no greater than the tendency to pick one of the other innocent persons, then 20% of those who had made a selection from the target-absent lineup should have chosen the previously seen innocent. If the actual selection rate were significantly greater than 20%, a transference effect would have been indicated (see Loftus, 1976). It should be noted that some investigators have taken the approach of comparing the rates at which transference errors are made by persons who have been subjected to a transference condition and by persons in a control

condition who have not been thus subjected. Clearly, a significantly greater rate of committing a transference error by the experimental group would be indicative of a transference effect. Either approach tests for the presence of a memory bias imposed by a transference condition.

At the lineup, eyewitnesses in Experiments 2 and 3 of Brown *et al.*'s (1977) study did indeed identify persons seen previously only in mugshots at significantly higher rates than they did previously unfamiliar foils. Interestingly, under Experiment 3's viewing conditions of only a brief glimpse of target persons, witnesses were as likely to indict a person on the basis of a single mugshot encounter as they were a target person encountered only once, that is a target person not also presented in mugshots.

Three other studies employing a transference design involving mugshot exposure have been published subsequent to that of Brown *et al.* (1977). Deffenbacher, Carr, and Leu (1981) tested whether a transference design would produce significant loss of memory for context of facial encounter at retention intervals of 2 min and 2 weeks. At both retention intervals, they found significantly higher false alarm rates to faces seen once previously in mugshots than to previously unfamiliar foils. More recently, Dysart, Lindsay, Hammond, and Dupuis (2001) provided a specific test of the transference effect and did not obtain a significant effect. Finally, Perfect and Harris (2003) tested for transference effects separately for young and older adults in three separate experiments, providing six different tests of whether a transference design produces negative effects on eyewitness memory. Significant transference effects were obtained in five of the six instances.

It should be noted that the transference effect is not a uniquely forensic phenomenon. In addition to the previously cited studies of eyewitness memory, memory researchers have for the past quarter of a century obtained strong support for a variety of transference-related phenomena. These phenomena include "familiarity without awareness" (Mandler (1980), "retention without awareness" (Roediger, 1990), the "false fame effect" (Jacoby, Kelley, Brown, & Jasechko, 1989), and failures in reality monitoring (Johnson & Raye, 1981) and source monitoring (Johnson, Hashtroudi, & Lindsay, 1993). The research literatures associated with each of these related implicit memory phenomena all converge upon the same conclusion: Dissociations between recognition and awareness of context are common.

There is yet another set of studies, not involving mugshot exposure, which nonetheless are relevant to an assessment of whether the phenomenon that has been labeled *unconscious transference* (Williams, 1963) can be reliably produced in studies employing transference designs. Unconscious transference is the term applied to the "transfer of one person's identity to that of another person from a different setting, time, or context" (Read *et al.*, 1990, p. 3). Williams (1963) coined the term in discussing an English murder case, which may have resulted in an innocent person being executed. It seems that one of the eyewitnesses who had identified the defendant had seen him briefly before the crime had occurred and may have engaged in an unconscious transference of identity from the actual perpetrator to the defendant.

Studies using transference designs wherein mugshots are not inserted between viewing of the target and a test of recognition memory, but which do include in that

test a face of a bystander who shared proximate temporal context with the target, would appear to be the laboratory analog of the case Williams (1963) has described. Loftus (1976) published the first such study and demonstrated a substantial transference effect. A posthoc one-tailed test performed on her data indicates that the obtained transference effect was not due to the inability of witnesses to discriminate between the target's face and other faces present during the original scenario; the target face was selected from a target-present lineup at a significantly greater rate than was the bystander face from a target-absent one. In testing for presence of the transference effect, Read *et al.* (1990) conducted five experiments and a total of six independent tests, each involving a TA lineup. Read *et al.*'s (1990) experiments were field experiments involving the witnessing of bystanders and targets by retail store clerks or by students in university classrooms going about their normal situation-appropriate activities. Tests of the difference between transference and control conditions in regard to the proportion of false positive identifications of the bystander resulted in two negative effects (control condition showing a higher proportion of transference errors), two effect sizes of zero, and two positive effects. Finally, Ross, Ceci, Dunning, and Tolia (1994) presented both transference and control condition witnesses with a film of a staged robbery. Transference witnesses saw an innocent male bystander in the film, whereas control witnesses viewed the same film without the innocent bystander. The former group of witnesses were nearly three times as likely to misidentify the bystander as the perpetrator of the robbery than were the control witnesses.

In addition to the passive effect of familiarity gained by mere mugshot exposure to a previously unfamiliar face, Gorenstein and Ellsworth (1980) sought to determine whether there was a further increase in interference with witness memory for the target face engendered by the witness actively choosing the mugshot of an innocent person.<sup>1</sup> Would witness commitment to a prior choice of a particular mugshot as the target person cause the mugshot image to be retained as strongly or even more strongly than that of the target? In brief, then, a mugshot commitment design involves a control group and an experimental group both being exposed to a target person, followed by mugshot exposure for the latter group. Witnesses in the experimental condition may be urged to pick an innocent person from the mugshots, despite the target's face not being included. At the lineup the target face may or may not be included, but the previously chosen mugshot image definitely is. Gorenstein and Ellsworth (1980) found that whereas control witnesses identified the target from a six-person photo array at a significantly above chance rate (.39), mugshot commitment witnesses failed to identify her at an above chance rate (.22) but did identify their mugshot choice's image at an above chance level (.44), a level somewhat greater than the rate at which control witnesses identified the target person.

Three additional studies employing a mugshot commitment design have been published since that of Gorenstein and Ellsworth (1980). Brigham and Cairns (1988) found the mugshot commitment manipulation both to decrease the identification

<sup>1</sup> Experimental psychologists have used the term *interference* to refer to the tendency for one activity to negatively impact memory for another activity. *Retrospective interference* would reflect a negative impact on memory for material encoded prior to the activity in question.

rate of the actual target (hit rate) and to increase the false alarm rate for the mugshot previously identified as the target. Dysart *et al.* (2001, Exp. 2), using only TA lineups likewise noted a substantial increase in the false alarm rate for witnesses who had made a prior commitment that a particular mugshot “is or maybe” the target. Most recently, Memon, Hope, Bartlett, and Bull (2002) in a slight modification of the mugshot commitment design, had all their witnesses take part in a TA lineup that contained a critical foil (shown prior in the mugshot array) and five new foils. Regardless of their particular choice, choosers of a mugshot from the mugshot array were significantly more likely to choose the critical foil from the TA lineup than were nonchoosers and the control group. This result suggests that the negative effects of mugshot exposure in a mugshot commitment design may not depend entirely on prior commitment to a particular critical foil.

Both the transference and mugshot commitment designs resemble the classic retroactive interference (RI) design used by experimental psychologists, in that both control and experimental conditions are exposed to targets, only the experimental condition is exposed to subsequent mugshots, and witnesses in both conditions are ultimately tested for their memory of the targets. However, both these designs differ from the classic RI design in that they include one or more of the previously seen mugshots in the final test of memory, whereas the RI design does not. The *pure RI design* would include only targets and new, previously unseen foils in a recognition test of memory such as a lineup. Thus, in the case of mugshot studies, the pure RI design would test whether exposure to intervening mugshots suppresses memory for the target items, per se, perhaps due to a form of memory blending in which the identities of a target image and that of a mugshot are confused at encoding (Perfect & Harris, 2003). The transference and mugshot commitment designs, on the other hand, permit assessment of decreases in memory for the target due to a failure of memory for source or context of encounter with a face actually previously seen in mugshots.<sup>2</sup>

Four studies employing the pure RI design with intervening mugshot exposure have been published in the past 25 years. Davies, Shepherd, and Ellis (1979) found a significant suppression of memory for targets when the intervening exposure to mugshots was conducted as a standard mugshot search. On the other hand, Cutler, Penrod, and O’Rourke (1986, Exp. 2), Cutler, Penrod, and Martens (1987), and Dysart *et al.* (2001, Exp. 1) either obtained decidedly smaller negative effects on witness memory or in the case of Cutler *et al.* (1987), a positive effect.

Given their widespread use in the furtherance of criminal investigations and their potentially biasing effect having been addressed by the courts (*e.g.*, *Simmons v. United States*, 1968), it has been important for the effects of mugshot exposure to witnesses to have been investigated. More than 25 years of research has accumulated concerning the possible biasing effects of mugshot exposure, research that has generated 32 independent tests of the effects of mugshot exposure on eyewitness memory. In that time, neither the seminal Shapiro and Penrod (1986) metaanalysis of the eyewitness literature nor any subsequent one has been concerned with the effect of mugshot exposure on eyewitness memory. Focused metaanalyses of other

<sup>2</sup>Deffenbacher *et al.* (1981) have referred to memory for the target face image(s) per se as item memory and memory for the source or context of encounter with a face as context memory.

eyewitness phenomena, such as weapon focus (Stebly, 1992), have been conducted after shorter intervals, 15 years in the case of weapon focus, and with as few as 19 relevant effect sizes (Stebly, 1992). Hence, it would be useful to have a firm estimate for both effect size and direction with regard to mugshot exposure, both with regard to proportion correct and the false alarm rate (the rate at which false positive identifications of lineup foils is made). This would allow us to test the hypothesis that mugshot exposure acts both to decrease the proportion of correct identification decisions as well as to increase the false alarm rate.

We also seek to identify variables that might moderate any consistent effect of mugshot exposure on accuracy of memory. In particular, we suspect that type of experimental design (pure RI, transference, or mugshot commitment design) may moderate the effect of mugshot exposure on eyewitness memory.

A final pair of related goals are to determine whether transference designs reliably result in increased rates of transference error and, if so, whether the inclusion of mugshots before the lineup produces any greater rate of transference error than not including mugshots. In other words, we seek to learn whether photobiasing lineups are any more productive of transference errors than are lineups resulting from transference situations not involving the potential bias of mugshot exposure, bystander situations, in particular.

## METHOD

### Sample

Given that the present review was part of a comprehensive metaanalysis project intended to update and extend the one conducted by Shapiro and Penrod (1986), a thorough search of social science citation retrieval systems was conducted. These systems included PsycINFO, Educational Resources Information Center (ERIC), *Sociological Abstracts*, *Dissertation Abstracts International*, *Dissertations on-line* (<http://www.contentville.com/content/dissertations.asp>), MEDLINE, and SOCIAL SCISEARCH (the *Social Science Citation Index*). These computer database searches were supplemented by more traditional search methods, including use of bibliographic citations in published research and in social science convention proceedings and contacting leading researchers, in order to identify the most recent published research.

No unpublished studies were included, inasmuch as the legal standards for proffered scientific testimony established by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993) have strengthened the preference by the legal system for metaanalytic conclusions based on a body of well conceived, well executed, and easily retrievable studies. The Court in *Daubert* enunciated four factors that judges should consider in evaluating the admissibility of expert scientific testimony. One factor to consider is whether the theory or technique has been subjected to peer review or publication. Thus, we are in agreement with Clark (2005), Deffenbacher, Bornstein, Penrod, and McGorty (2004), and Reisberg and Heuer (2006) that to survive a *Daubert* review, conclusions about

effect size and direction should be based on published studies. We likewise agree with the point made by Clark (2005) that without peer review, unpublished results or the methodological details of the studies which produced them, are not readily available for necessary public scrutiny and evaluation.

Now, certainly the peer review process may have the effect of keeping zero or nearly zero effect sizes, which are more likely to characterize unpublished studies, from accumulating in the scientific literature. However, many unpublished studies have earned their status by having one or more of the following characteristics, low power, ineffective manipulations, severe confounds, or the need for one or more additional experiments to provide adequate controls or to rule out alternative explanations of results. That is, unpublished studies have often either not passed the peer review process or have not been subjected to it, because the investigator has already recognized the problem(s) with the study. There is yet another problem with trying to include unpublished studies. Unlike the situation for published studies, there is no systematic means for gathering unpublished studies. Consequently, one ends up with a set of unpublished papers that is likely not representative of all the extant unpublished papers. The unpublished papers obtained include those one happens to know about, from investigators one happens to know (who have perhaps previously published on the topic—though others might not have), and who happen to have reasonably clean written drafts of their unpublished results and data files. Nevertheless, we shall present a posthoc analysis in the discussion section of this paper, showing that even the weakest of our overall metaanalytic conclusions would not have been changed had we been able to add four additional effect sizes from unpublished studies (Stebly, 1992), in the present case, the exact same effect sizes as those produced in the four unpublished studies in Stebly's sample of 19 studies. These four effect sizes have a mean weighted by sample size of  $-.09$ .

There were two final study samples. The first sample comprised 11 published articles, providing 32 independent tests of the hypothesis that prior mugshot exposure decreases accuracy on a subsequent lineup. This sample included work published between 1977 and 2003, work yielding a total of 1664 participants.<sup>3</sup> Sample sizes across the 32 tests of the mugshot exposure effect ranged from 26 to 276 ( $M = 79.24$ ).

In order to test whether a reliable unconscious transference effect exists, another sample of seven published articles (1976–2003) was collected, four of which were included in the prior sample, inasmuch as in these prior articles, investigators used a transference design to test for the presence of a mugshot exposure effect. Investigators in the other three articles employed transference designs that did not include mugshot exposure, bystander studies. The entire sample yielded 19 independent tests of the transference effect and a total of 1145 participants, with sample sizes ranging from 24 to 165 ( $M = 60.26$ ).

<sup>3</sup> Even though the cutoff year for the aforementioned comprehensive metaanalysis project was 2002, the Perfect and Harris (2003) article was included for the limited purpose of conducting the present metaanalysis focused just on mugshot exposure effects. This study is one of three published since 2000 and adds six independent tests to a prior total of 26 tests of the hypothesis that mugshot exposure is detrimental to eyewitness memory accuracy. The initial year for the current narrowly focused metaanalyses was determined by the fact that the first published study of mugshot exposure effects occurred in 1977.

Finally, it should be noted that even though the current metaanalyses were intended to update and extend the Shapiro and Penrod (1986) metaanalysis, the current sample of studies has little overlap with theirs. The set of studies included in their comprehensive metaanalysis contains only two of 14 articles included in the current sample. These two articles contributed only three of the 32 effect sizes involved in our first two metaanalyses.

### Study Characteristics

The previously mentioned comprehensive metaanalysis project, from which the current study springs, has involved the coding of several dozen independent variables for each of 469 existing studies of face recognition from both the eyewitness identification and laboratory face recognition memory traditions. A prior study of interrater reliability at coding 80 of these independent variables showed an average agreement of 93%. Coded variables included stable (*e.g.*, sex and race) and malleable (*e.g.*, disguise) characteristics of both witnesses and targets as well as situational (*e.g.*, target exposure duration) and procedural (*e.g.*, lineup presentation) factors. Variables worthy of mention for their potential usefulness in the present review included type of study (eyewitness identification study or laboratory face recognition study), whether or not a study employed a staged crime, whether the lineup included the target (TP) or not (TA), number of participants, age of participants, number of mugshots inserted between target exposure and subsequent memory test, and obviously, whether mugshot exposure was manipulated. Dependent variables recorded were proportion correct, hit and false alarm rates, if provided, for TP lineups, correct rejection and false alarm rates, when provided for TA lineups, and the signal detection measures,  $d'$  and  $\beta$ .

### Statistics

To test the statistical reliability of an estimate of the typical effect size found in any particular metaanalysis, we have adopted the Stouffer method (Rosenthal, 1995). Here a metaanalytic  $Z$  ( $Z_{ma}$ ) was calculated by combining  $Z$ -scores associated with individual tests of the hypotheses that mugshot exposure and transference designs negatively impact eyewitness memory. The resulting algebraic sum, when divided by  $\sqrt{k}$ , where  $k$  is the number of independent estimates of the effect size, yields the metaanalytic  $Z$ . The probability associated with the metaanalytic  $Z$  is the overall probability of a Type I error associated with the observed pattern of results. Inasmuch as  $Z$  provides an unweighted estimate of the overall probability level, a metaanalytic  $Z$  ( $Z_{mn}$ ) was also calculated which weighted individual  $Z$ -scores by the sample size of the study. This allows estimation of population parameters with greater emphasis on larger samples and their more reliable parameter estimates.

It should be noted that whenever recovery of sample sizes and either proportion correct (hits, correct rejections) identifications or false alarm rate per condition permitted, the  $Z$ -score entered into the meta-analysis was one calculated for the dif-

ference between proportions. When an exact  $Z$ -score could not be calculated for a given effect size estimate, a  $Z$ -score associated with the  $p$  value for the estimate was entered, 1.65 for  $p = .05$ , for instance. When a test of the hypothesis was reported as not significant, but no statistics were cited, the conservative procedure of entering  $Z = 0.00$  was followed (Rosenthal, 1995).

All studywise differences between proportions for experimental (mugshot exposure or transference design) and control conditions (no mugshot exposure or no transference design) were converted into the effect size measure  $h$ , inasmuch as  $h$  is the coefficient recommended by Cohen (1988) for testing for differences between proportions. Hence, mean effect size for any set of studies and its associated 95% confidence interval is expressed in terms of  $h$ .<sup>4</sup>

Finally, given that we have included only published studies in the present metaanalysis, it is quite clear that our sample of studies is not a random sample of all studies that have ever been conducted seeking to test the effects of mugshot exposure. Rosenthal (1995) has pointed out that it is quite likely that published studies have reported lower probabilities of a Type I error than have those studies "squirreled away in file drawers." The concern in regard to this "file drawer problem" is that a sufficient number of studies averaging null results could threaten a metaanalytic conclusion. We have, therefore, adopted Rosenthal's (1995) suggested procedure for calculating a failsafe  $N$  ( $N_{fs}$ ) in order to determine the number of unknown or not retrieved studies averaging null results required to increase the probability of a Type I error to the just significant level of  $p = .05$ . Actually, inasmuch as this number is usually a whole number plus a fractional number of studies, we have adopted the rule of rounding to the next higher number. Thus, most values of the failsafe  $N$  that we report in connection with a metaanalytic  $Z$ , represent the number of additional null results studies required to increase the probability of a Type I error to a value slightly greater than .05. Clearly, the failsafe  $N$  represents a "tolerance for future null results" (Rosenthal, 1995). We agree with Deffenbacher *et al.* (2004) that at an absolute minimum the failsafe  $N$  must be at least as large as the number of independent estimates of effect size that went into calculating the metaanalytic  $Z$ .

## RESULTS

### Metaanalysis 1: Mugshot Exposure and Proportion Correct

#### *All Tests*

Our first task was to determine the overall status of the hypothesis that mugshot exposure decreases proportion correct (hits, correct rejections) on tests

<sup>4</sup>Inasmuch as the power to detect a difference between proportions is not simply a function of the magnitude of the difference, but varies as a function of where a given difference occurs along the scale of differences between zero and one, an arcsin transformation (nonlinear) of these differences has been suggested by Cohen (1988). The statistic  $h$  is the resulting transform of a given difference in proportions. For instance, the following pairs of proportions all yield approximate values of  $h = .20$ : .05 and .10, .20 and .29, .40 and .50, .60 and .70, .80 and .87, and .90 and .95.

of eyewitness memory for faces. For this analysis, proportion correct for the control condition was subtracted in each instance from that for the experimental condition. Hence, a negative metaanalytic  $Z$  would indicate a debilitating effect of mugshot exposure on proportion correct performance. The overall proportion of correct identifications at a lineup for the mugshot exposure conditions was .43; the comparable figure for the control conditions was .50 correct. Mean effect size,  $h$ , for this analysis was  $-.15$  (95%  $CI$ :  $-.33$  to  $+.03$ ); median effect size was  $-.17$ . The unweighted metaanalytic  $Z_{ma}$  was  $-2.98$ ,  $p < .01$ ,  $N_{fs} = 35$ . Weighting each of the 15 independent tests of the hypothesis by sample size yielded  $Z_{mn} = -2.11$ ,  $p < .025$ ,  $N_{fs} = 10$ , a failsafe  $N$  that is less than our previously stated standard. Nevertheless, taken together, these analyses provide at least some support for the hypothesis that mugshot exposure inserted between exposure to the target and a subsequent lineup test has a negative effect on proportion correct measures of eyewitness memory.

We next applied a test of the homogeneity of variances across the sample of weighted effect sizes to determine whether the degree of variability exceed that expected on the basis of sampling error alone. A  $\chi^2$  value of 84.57, with  $df = 14$ ,  $p < .001$ , indicated that the degree of variability did indeed exceed that expected on the basis of sampling error. In subsequent analyses, we tested for moderator variables which might account for the considerable variability in the effect sizes for individual tests of the hypothesis (effect size range:  $-.694$  to  $+.447$ ,  $s = .32$ ; see Table 1).

Given the relatively small sample of effect sizes and their considerable variability, we were concerned about two potential problems in interpreting results of our moderator analyses. If the variance in effect sizes were "carved up" in several overlapping ways, multicollinearity could be a problem. Further, where there were only a few effect sizes of considerable variability associated with a given moderator variable, it would be possible for one or two large effect sizes to produce a significant metaanalytic  $Z$ , but for the 95%  $CI$  to include zero.

To mitigate these concerns, we took two decisions. First, we decided that we would limit each moderator analysis to two moderator variables, two variables most directly germane to studies of mugshot exposure, the number of mugshots exposed and the nature of the experimental design for producing and testing for effects of mugshot exposure. That is, we limited ourselves to carving the effect size variance in only two ways. Further, if analysis were to reveal a clear confounding of the two moderator variables, we decided that we would then report only the results for one, the one producing the more trustworthy effect size. We defined a *trustworthy effect* for present purposes as one that was associated with a statistically significant metaanalytic  $Z$ , a 95%  $CI$  that did not include zero, and a failsafe  $N$  that was at least twice as great as the number of effect sizes in the sample or subsample.

### *Moderator Analyses*

Inasmuch as four of the 15 tests of the hypothesis that mugshot exposure decreases proportion correct were studies wherein the target(s) appeared in mugshots as well as in the final test of memory (Brown *et al.*, 1977, Exp. 2–3; Lindsay, Nosworthy, Martin, & Martynuck, 1994, Exp. 1, both target present and target absent lineups), one might therefore predict not a negative effect of mugshots in

**Table 1** Identification Accuracy Effect Sizes (Studies Ordered by Date)

Author	<i>N</i>	Proportion correct ( <i>h</i> )	False alarms ( <i>h</i> )
Loftus 1976 <sup>a</sup>	25		1.26
Brown et al., 1977, Exp. 2	64	.28	.35
Brown et al., 1977, Exp. 3	146	.45	.26
Davies et al., 1979	60	-.45	
Gorenstein & Ellsworth, 1980	36	-.37	.60
Deffenbacher et al., 1981, 2 min	27	-.44	.77
Deffenbacher et al., 1981, 2 week	26	-.24	.41
Cutler et al., 1986, Exp. 2	276	-.17	
Cutler et al., 1987	165	.21	
Brigham & Cairns, 1988	99	-.69	.94
Read et al., 1990, Exp. 1 <sup>a</sup>	31		.00
Read et al., 1990, Exp. 2 <sup>a</sup>	20		-.61
Read et al., 1990, Exp. 3 <sup>a</sup>	69		-.43
Read et al., 1990, Exp. 4 <sup>a</sup>	49		.00
Read et al., 1990, Exp. 5A <sup>a</sup>	96		.14
Read et al., 1990, Exp. 5B <sup>a</sup>	131		.34
Lindsay et al., 1994, Exp. 1, TP	69	.00	.00
Lindsay et al., 1994, Exp. 1, TA	70	.00	.00
Ross et al., 1994, Exp. 1 <sup>a</sup>	165		.82
Dysart et al., 2001, Exp. 1	104	-.08	
Dysart et al., 2001, Exp. 2 <sup>b</sup>	88	-.37	.87
Dysart et al., 2001, Exp. 2 <sup>c</sup>	88	.04	-.32
Memon et al., 2002	112	-.46	.46
Perfect & Harris, 2003, Exp. 1A <sup>d</sup>	30		.07
Perfect & Harris, 2003, Exp. 1B <sup>e</sup>	51		.74
Perfect & Harris, 2003, Exp. 2A <sup>d</sup>	45		.64
Perfect & Harris, 2003, Exp. 2B <sup>e</sup>	48		.35
Perfect & Harris, 2003, Exp. 3A <sup>d</sup>	30		.58
Perfect & Harris, 2003, Exp. 3B <sup>e</sup>	30		1.44

<sup>a</sup>Transference studies without mugshot exposure.

<sup>b</sup>Mugshot commitment design.

<sup>c</sup>Transference design with mugshot exposure.

<sup>d</sup>Young adult participants.

<sup>e</sup>Older adult participants.

these instances, but rather no effect or perhaps a positive effect. That is, having an extra exposure of the target under ideal viewing conditions might well increase the accuracy of eyewitness memory on a subsequent test, rather than decrease it. Such an effect would impose another sort of bias on any subsequent lineup, especially if the eyewitness had not identified the actual perpetrator from the mugshot presentation. If the perpetrator's image being present in the mugshots increased the probability that he would subsequently be identified at the lineup, justice would have been served, but one of the assumptions of the law concerning eyewitness testimony would not have been met; the perpetrator would have been identified not based strictly on the eyewitness's uncontaminated memory of the crime scene.

The proportion correct for the mugshot exposure conditions, taken across all four aforementioned studies, was .55 versus a figure of .48 for the control conditions. The metaanalytic  $Z$  was statistically reliable,  $Z_{ma} = 2.33$ ,  $p < .01$ ,  $N_{fs} = 4$ . Inasmuch as the confidence interval associated with the average effect size of  $h = +.18$  included zero (95% *CI*:  $-.17$  to  $+.53$ ), and the failsafe  $N$  did not meet our criterion for a trustworthy effect, it would be safest to conclude that there is no biasing effect on proportion correct induced by including the target(s) in mugshots as well as a subsequent lineup.

Desiring a measure of possibly biasing effects of mugshot exposure, without the contamination of target presence in the mugshots, a moderator analysis was conducted on the 11 tests of the mugshot exposure hypothesis in which targets did not subsequently appear in mugshots. Overall proportion correct for the mugshot exposure conditions in this instance was .48, with the comparable figure for the control conditions being .61,  $Z_{ma} = 4.91$ ,  $p < .0001$ ,  $N_{fs} = 87$ . This uncontaminated test of possible mugshot exposure bias yielded an effect size nearly twice as strong as that produced by the preceding omnibus test,  $h = -.28$  (95% *CI*:  $-.11$  to  $-.45$ ).

Further moderator analyses were confined to just these 11 effect sizes. Applying the moderator variable of experimental design produced three subsets of effect sizes, four associated with studies employing a *pure RI design*, four associated with studies using a *mugshot commitment design*, and three associated with studies using a *transference design*. In the first instance, the metaanalytic  $Z$  was not significant,  $Z_{ma} = -1.52$ ,  $p < .10$ . Mean effect size was  $h = -.12$ , 95% *CI*:  $-.55$  to  $+.31$ . Thus, the amount of retroactive interference with memory for the target generated simply by interpolation of mugshots does not appear to reduce significantly accuracy (proportion correct) at a subsequent lineup. Identification of a mugshot image as that of the target and the subsequent appearance of that image or person at the lineup is a different matter, however. For the mugshot commitment design, the metaanalytic  $Z_{ma} = -4.33$ ,  $p < .0001$ ,  $N_{fs} = 24$ , with an associated effect size  $h = -.48$  (95% *CI*:  $-.24$  to  $-.72$ ). Hence, use of a mugshot commitment design produces a sizeable negative impact upon lineup accuracy, with mean proportion correct of .32 for the mugshot commitment condition versus .55 for the control condition. Apparently identification from mugshots of either a foil or innocent person as the target, interferes sufficiently with a witness's memory representation of the actual target so as to reduce both the hit rate and correct rejection rate for target present and target absent lineups, respectively. Finally, for the transference design, overall proportion correct for the transference condition was .66 and .76 for the control condition,  $Z_{ma} = -2.63$ ,  $p < .01$ ,  $N_{fs} = 5$ . The corresponding effect size,  $h$ , was  $-.21$ , with a rather broad 95% *CI* of  $-.81$  to  $+.39$ . All in all, the negative impact on proportion correct imposed by a transference design does not appear to be a very trustworthy effect.

A careful analysis of the 11 effect sizes from studies that did not include the target(s) in mugshots showed no discernable relationship between type of experimental design employed and the number of mugshots to which witnesses were exposed, whether a relatively large number or a relatively small one. Hence, with just two different "cuts" of this set of effect sizes, there would appear to be no problem of multicollinearity affecting interpretation of results.

The *number of mugshots* interpolated between exposure to the target and subsequent lineup test was therefore tested as a candidate for moderating the effect of mugshot exposure on eyewitness memory performance. For the 15 tests of the hypothesis that mugshot exposure decreases lineup hit or correct rejection rate, the number of interpolated mugshots ranged from 12 to 600, with a median of 21. In the case of the six studies with above the median numbers of mugshots (41 or more) and which did not present targets in the mugshots,  $Z_{ma} = -1.86$ ,  $p < .05$ ,  $N_{fs} = 2$ .

The corresponding effect size,  $h$ , was  $-.14$  (95% *CI*:  $-.40$  to  $+.12$ ). Given such a low failsafe  $N$  and a confidence interval that includes zero, it is probably best to conclude that a relatively high number of interpolated mugshots is not associated with a deleterious effect on proportion correct at a lineup. Interestingly enough, however, for the five studies with below the median numbers of mugshots and which excluded targets from them, the metaanalytic  $Z$  and corresponding effect size were considerably larger,  $Z_{ma} = -5.23$ ,  $p < .0001$ ,  $N_{fs} = 46$ ,  $h = -.44$  (95% *CI*:  $-.24$  to  $-.64$ ).

## Metaanalysis 2: Mugshot Exposure and the False Alarm Rate

### *All Tests*

We first tested the hypothesis that mugshot exposure decreases accuracy at a subsequent lineup by virtue of increasing the false alarm rate. Clearly, unlike Metaanalysis 1, analysis of whether or not the target appeared among the interpolated mugshots was irrelevant. Parallel to our analysis of the proportion correct measure, we subtracted the false alarm rate for the control condition in each instance from that for the experimental condition. Thus, a positive metaanalytic  $Z$  would indicate a debilitating effect of mugshot exposure on the false alarm rate, that is to say, by increasing it. The overall false alarm rate at a lineup for mugshot exposure conditions was  $.37$ , and the comparable figure for the control conditions was  $.15$ , less than half the former rate. The unweighted metaanalytic  $Z_{ma} = 9.94$ ,  $p < .0001$ ,  $N_{fs} = 602$ . Weighting each of the 17 tests by sample size yielded  $Z_{mn} = 9.05$ ,  $p < .0001$ ,  $N_{fs} = 499$ . Mean effect size,  $h$ , for the unweighted analysis was  $+.48$ , with a 95% confidence interval extending from  $+.26$  to  $+.70$ . Median effect size was  $+.41$ . These analyses support the hypothesis that mugshot exposure increases the false alarm rate at a subsequent lineup.

Seeking to determine whether the variability in effect sizes exceeded that expected from sampling error alone, we applied a test for homogeneity of variance to the sample of weighted effect sizes;  $\chi^2 = 36.42$ , with  $df = 16$ ,  $p < .01$ . Subsequently, we tested for moderator variables which might account for the substantial variability in effect sizes for individual tests of this hypothesis (effect size range =  $-.318$  to  $+1.442$ ,  $s = .42$ ; see Table 1).

### *Moderator Analyses*

Inasmuch as the four studies employing a *pure RI design* only reported overall proportion correct, and false alarm rates for target present lineups could not be inferred due to the false rejection rate not being reported, variability in effect sizes moderated by a pure retroactive interference design could not be ascertained. However, four studies using a *mugshot commitment design* did provide tests of whether mugshot exposure in the context of that particular design increases the false alarm rate at a subsequent lineup. There was indeed support for a substantial moderating effect for application of this design, with a mean false alarm rate for the mugshot commitment condition of  $.53$  versus just  $.20$  for the control condition,

$Z_{\text{ma}} = +6.09, p < .0001, N_{\text{fs}} = 51$ . The average effect size,  $h$ , was  $+.72$  (95% *CI*:  $+.38$  to  $+1.06$ ). Finally, 11 studies combining mugshot exposure within the context of a *transference design* produced evidence of a significant moderating effect of this particular design on the lineup false alarm rate (.38 for the transference condition as compared with .16 for the control condition). The unweighted metaanalytic  $Z$  was  $+8.66, p < .0001$ , and the failsafe  $N$  was 295 additional studies averaging null results. The corresponding average effect size,  $h$ , was  $+.44$ , the 95% confidence interval for which stretched from  $+.11$  to  $+.77$ . It is fair to conclude here that the effect size for the mugshot commitment design was decidedly larger than that for the transference design, some 65% larger.

A careful analysis of the sample of effect sizes again showed no apparent confounding of type of experimental design and the number of mugshots to which witnesses had been exposed. Therefore, we again decided to test for the moderating influence of the relative *number of mugshots* exposed.

In this instance, the number of interpolated mugshots above the median ranged from 16 to 600, and the number below the median ranged from 8 to 15. As occurred with the moderating effect of number of mugshots on proportion correct, exposure of a below-median number of mugshots (eight studies) in this case had a greater impact on the false alarm rate than did exposure to an above-median number of them (nine studies). In the former case,  $Z_{\text{ma}} = 7.66, p < .0001, N_{\text{fs}} = 166$ ; mean  $h = +.44$  (95% *CI*:  $+.26$  to  $+.62$ ). In the latter case,  $Z_{\text{ma}} = 6.42, p < .0001, N_{\text{fs}} = 128$ ; mean  $h = +.35$  (95% *CI*:  $-.18$  to  $+.88$ ). Not only is the effect size for a relatively small number of mugshots exposed some 25% larger than that for a relatively large number of mugshots exposed, but it is a more trustworthy result.

### Metaanalysis 3: Transference Errors

#### *All Tests*

We began by testing the hypothesis that irrespective of whether mugshots are exposed to eyewitnesses between initial exposure to the target person and subsequent identification test, use of a transference design will increase the rate at which transference errors are made. Here again, parallel to our metaanalysis of false alarm responses generated as a function of mugshot exposure, we subtracted the false alarm rate (to faces previously seen during mugshot exposure) for the control condition in each instance from that for the transference condition. Hence a positive metaanalytic  $Z$  would indicate a debilitating effect of transference conditions upon identification accuracy. That is to say, increased likelihood of a transference error decreases the likelihood of a correct identification at a lineup. When a transference condition had been created, the overall rate at which transference errors were made across our sample of 19 independent tests of the transference hypothesis was .33. The comparable figure for the control condition was .16, less than half the rate for the transference condition.

The unweighted metaanalytic  $Z$  was  $9.24, p < .0001, N_{\text{fs}} = 581$  studies. Weighting each of the 19 tests by sample size yielded a metaanalytic  $Z$  of  $8.89, p < .0001$ , with a failsafe  $N$  of 537 studies. Mean effect size,  $h$ , for the unweighted

analysis was  $+0.34$ , with a confidence interval extending from  $+0.08$  to  $+0.60$ . Median effect size was  $+0.35$ . These analyses support the hypothesis that whether exposed to mugshots or not, presence of a transference condition increases the particular sort of false alarm error known as a transference error.

Applying a test of homogeneity of variance to the sample of weighted effect sizes,  $\chi^2 = 64.58$ , with  $df = 18$ ,  $p < .001$ . As was the case with the two previous metaanalyses, the degree of variability exceeded that which would be expected on the basis of sampling error alone. Hence, we sought to determine whether type of transference design and age of the eyewitness could account for some of this variability in effect size.

### *Moderator Analyses*

Considering just the eight transference design studies we have labeled as bystander studies,  $Z_{ma} = 4.08$ ,  $p < .0001$ ,  $N_{fs} = 42$ . The corresponding mean effect size,  $h$ , was  $+0.19$  (95% *CI*:  $-0.33$  to  $+0.71$ ). For the 11 transference designs which included mugshot exposure, the metaanalytic  $Z$  was  $8.66$ ,  $p < .0001$ , with a failsafe  $N$  of 295 studies. In this latter instance, the average effect size,  $h$ , was  $+0.44$ , with a 95% confidence interval which did not include zero, extending from  $+0.11$  to  $+0.77$ . Hence, though statistically reliable in both cases, transference conditions created as a function of witnesses viewing someone at a lineup who had also been seen in a prior mugshot were more potent in producing transference errors (mean transference error rate of  $.38$  versus  $.16$  for the control condition) than were transference conditions created by the viewing at a lineup of a previously seen bystander (transference error rates of  $.27$  and  $.17$ , respectively). The former sort of transference design is associated with an effect size more than twice that of the latter.

The second moderator analysis was performed only on the subset of 11 effect sizes associated with transference designs in which witnesses had been exposed to mugshots. This particular sample of effect sizes was chosen, inasmuch as Perfect and Harris (2003) performed separate tests of the transference hypothesis on younger (mean ages of 20–22) and older adults (mean ages of 63–71) in three different experiments, in all of which the witnesses had been exposed to a set of mugshots interpolated between encounter of the targets and subsequent lineup tests. In none of the bystander studies had witness age been employed as an independent variable. Hence, the moderator analysis was performed on a sample of three effect sizes associated with older adult witnesses and eight effect sizes associated with younger adults.

Perhaps not surprisingly, the transference effect was statistically significant for both age groups. For younger adults,  $Z_{ma} = 4.61$ ,  $p < .0001$ ,  $N_{fs} = 55$ , and for older adults,  $Z_{ma} = 9.08$ ,  $p < .0001$ ,  $N_{fs} = 89$ . However, older witnesses were much more likely to commit a transference error, with a mean transference error rate of  $.57$  as compared with a rate of  $.30$  for younger adults (see Bartlett & Fulton, 1991); mean effect size,  $h = +0.85$  (95% *CI*:  $-0.52$  to  $+2.22$ ) for the older adults. The mean effect size for younger witnesses,  $h$ , was  $+0.35$  (95% *CI*:  $+0.06$  to  $+0.64$ ). Given the very small number of tests of the transference hypothesis with older witnesses, and the resulting very large confidence interval, this moderator analysis result should be interpreted with caution.

## DISCUSSION

### All Tests: Proportion Correct and False Alarms

Even though exposure of eyewitnesses to mugshots may well be a valuable crime investigative tool, we have adduced considerable evidence for the hypothesis that mugshot exposure decreases accuracy for a given witness at a subsequent lineup, both in terms of decreased likelihood of being correct (including hit and correct rejection rates) and increased probability of committing a false alarm. When we excluded those studies which included targets in a set of mugshots as well as at the lineup, we noted that exposure to mugshots decreased proportion correct at the lineup, the mean effect size being  $-.28$ , with a 95% confidence interval not including zero. The metaanalysis of false alarm rates produced an average effect size of  $+.48$ , with the lower bound of the confidence interval being  $+.26$ . However, evidence of the deleterious effect of mugshot exposure on proportion correct at the lineup is on somewhat less solid ground than is the evidence of mugshot exposure elevating the lineup false alarm rate. In the former case, the failsafe  $N$  associated with the unweighted metaanalytic  $Z$  requires only 87 additional studies averaging null results to increase the probability of a Type I error to the just significant level of  $p = .05$ ; in the latter case, the failsafe  $N$  is 602 studies. Thus, the effect of mugshot exposure in increasing the false alarm rate is both 70% larger and has greater long-run trustworthiness than its effect in decreasing the proportion correct at subsequent lineups.

### Moderator Variables

In addition to establishing the direction and magnitude of the effect of mugshot exposure on proportion correct and the false alarm rate at a subsequent lineup, we had sought to identify variables that might moderate any consistent effects of mugshots on the fidelity of eyewitness memory measurement. In the case of both dependent measures, we identified two different moderators of the effect of mugshot exposure, type of experimental design, and number of mugshots exposed.

The particular method by which mugshot exposure's effects were tested produced strong differences in effect size on both proportion correct and false alarm rate. Mugshot commitment designs yielded the largest effect sizes in both instances,  $-.48$  for proportion correct and  $+.72$  for the false alarm measure. Yielding smaller deleterious effects on memory performance were transference designs, producing an average effect size of  $-.21$  for proportion correct and  $+.44$  for false alarms. The moderating effect of pure retroactive interference designs could only be determined for the proportion correct measure, resulting in an effect size of just  $-.12$ .

It is interesting that mugshot-induced retroactive interference, in its purest form, has only a very limited negative impact upon eyewitness memory. As noted earlier, the pure retroactive interference design would include only previously unseen foils at a lineup. Mugshot commitment and transference designs, on the other hand, both include one or more faces in the ultimate test of memory that had been included in the mugshots. In our sample of studies, the photobias nature of the

lineup in a transference design increased the negative effect size for proportion correct by 75%, from  $-.12$  (effect size for pure RI designs) to  $-.21$ . The mugshot commitment design compounds this photobias by including in the lineup not only a person whose face was previously included among the mugshots but also by the previously viewed person having been identified by the witness *as* the target person. If the witness had indeed been willing to identify a previously unseen mugshot as the face of the target person, then clearly the witness must not have had a very distinctive encoding of the target's face, and it would neither be surprising that there would be a lowered probability of a correct identification response nor an increased probability of a false alarm in a subsequent lineup. For our study sample, use of a mugshot commitment design increased the negative effect size for proportion correct by a factor of four, from  $-.12$  to  $-.48$ . Use of a mugshot commitment design also increased the effect size for the false alarm rate by approximately 65%, from an effect size of  $+.44$  for transference designs to an effect size of  $+.72$ .

Thus, we have been able to provide substantial quantitative support for the empirical generalization that exposure to a mugshot of a suspect increases the probability that the eyewitness will subsequently choose that suspect in a lineup, especially when the real suspect is absent. Not only is the photobias resulting from the presence of transference conditions productive of such a bias, but identification of a mugshot *as* the suspect (mugshot commitment) substantially furthers the bias. Now Kassin, Tubb, Hosch, and Memon (2001) have reported in the most recent survey of eyewitness testimony experts ( $N = 64$ ), that 79% of the experts believed mugshot-induced bias to be either a "generally reliable" or "very reliable" empirical phenomenon. Similarly, Wise and Safer (2004) in their survey of 160 US judges as to their knowledge and beliefs about eyewitness testimony have indicated that fully 78% of them believed the mugshot-induced-bias phenomenon to be "generally true." Given our results, there would seem little reason that these endorsement rates for both experts and judges should not be even higher in the future. In addition, our results make it clear that the likelihood of mugshot-induced bias is relatively low, unless a person has appeared both in mugshots and at a subsequent lineup.

Interestingly, exposure to relatively low numbers of mugshots had a more negative effect on witness memory than did greater numbers of mugshots. Perhaps it is easier to encode for memory storage a small set of mugshots in a reasonably distinctive fashion than it is for a much larger set. Certainly more distinctively encoded items are more easily retrieved, as they might be later at a lineup test of memory. In addition to a task difficulty explanation of this phenomenon, there is one that is related to motivation or attention: Perhaps exposure to a relatively large number of mugshots motivates witnesses to put minimal effort into encoding the photos, because the police clearly have no real suspect, and the witnesses are just searching. On the other hand, with a relatively small array of photos, witnesses might believe that they are in something like a focused lineup situation in which the police have a suspect; hence, witnesses might well be motivated to pay greater attention to and to engage in more distinctive encoding of the photos, including that of the person who may subsequently appear in a lineup.

## Transference Errors

Our final pair of related goals were to determine whether transference designs are reliably productive of increased rates of transference error and if this be the case, whether mugshot exposure is any more productive of transference error than are transference situations not involving mugshot exposure. Answering the latter question will determine whether transference conditions enabled by the criminal justice system, ones including photobiasing lineups, result in greater rates of transference error than do transference conditions that are more nearly “accidents of nature.” Considered in their entirety, use of transference designs increased the likelihood of transference errors, with an associated effect size of  $+0.34$  and a confidence interval not including zero. Transference designs involving mugshot exposure were associated with an effect size ( $+0.44$ ) more than twice that of transference designs having no possibility of a subsequent photobiasing lineup ( $+0.19$ ). Perhaps the smaller effect size for innocent bystanders observed at the crime scene obtains for the same reason as the earlier cited false fame effect (Jacoby *et al.*, 1989). The latter effect requires that the stimulus be perceived just well enough to be encoded but not so well encoded that the witness recalls the source. On the other hand, a witness viewing at least a relatively small number of mugshots shortly after the crime likely inspects them with a greater degree of thoroughness, given that one of them may be an image of the perpetrator. Again, the latter more distinctively encoded images are more likely to be mistakenly retrieved at a later target-absent lineup. At any event, both questions we have posed concerning the nature of transference effects can be answered in the affirmative.

Even though Wise and Safer’s (2004) survey did not ascertain what their sample of US judges believed to be true with respect to the phenomenon of unconscious transference, Kassin *et al.* (2001) did find that just 60% of experts believed unconscious transference to be generally reliable or very reliable. We should note that Kassin *et al.* defined unconscious transference for their sample of experts as, “Eyewitnesses sometimes identify as a culprit someone they have seen in another situation or context” (p. 408). From the present vantage point, this statement is remarkably similar to Kassin *et al.*’s definition of mugshot-induced bias, “Exposure to mug shots of a suspect increases the likelihood that the witness will later choose that suspect in a lineup” (p. 408). It is not clear whether very many of Kassin *et al.*’s experts had this similarity in mind when they rated the reliability of the phenomenon of unconscious transference. Given the absence of a relevant metaanalysis, if the experts had been relying only on the number of significant results favoring the transference hypothesis, three out of eight tests in the literature that had been available to them, then the 60% endorsement figure would appear to be unrealistically high. Perhaps, then, some of the experts recognized that mugshot-induced bias is oftentimes just another instance of the transference effect.

Given the results of our metaanalysis of the results of transference studies, the question becomes one of how an expert should now view the reliability of the transference phenomenon. Even the effect size ( $+0.19$ ) associated with the bystander studies also is associated with a metaanalytic  $Z$  ( $p < .0001$ ) that is sufficiently

reliable that the failsafe  $N = 42$ . That is, 42 additional undiscovered studies or studies to be conducted in the future, averaging null results, would be required to increase the probability of a Type I error to the just significant level of  $p = .05$ . If one were to make a reliability judgment based on just these studies, perhaps an endorsement rate of 60% of experts viewing transference to be a generally reliable phenomenon would not be unreasonable. However, if one properly considers the entire corpus of transference studies as tests of the hypothesis that transference conditions will increase the rate at which transference errors are made, then our results would not make unreasonable an endorsement rate of “generally reliable or very reliable” approaching 100%. After all, the effect size for the entire corpus of transference studies was  $+.34$ , with a confidence interval not including zero. Furthermore, the metaanalytic  $Z$  was associated with a failsafe  $N$  of nearly 600 studies.

### Limitations

Two limitations might be proposed to the generality of findings produced by our metaanalyses. First, it might be objected that we should have included as many unpublished studies of the effects of mugshot exposure as we could have uncovered. We shall not review here arguments for including only published studies, arguments made in the method section of this paper. However, as promised, we will present a posthoc analysis demonstrating that even the weakest of our overall metaanalytic findings would not have been changed had we included effect sizes from as many as four additional unpublished studies. For the sake of argument, had we included precisely the same effect sizes as those produced by Steblay's (1992) set of four unpublished studies (15 published effect sizes), our findings regarding the effect of mugshot exposure on proportion correct at a subsequent lineup would have changed very little. Mean effect size,  $h$ , would have decreased from  $-.15$  to  $-.12$ , with the confidence interval being almost the same, actually a bit tighter,  $-.27$  to  $+.03$ . Median effect size would have remained the same at  $-.17$ . The metaanalytic  $Z_{ma} = -2.72$ ,  $p < .01$ , essentially the same values as we have already reported, and the failsafe  $N$  would have decreased by only one study, from 35 to 34 undiscovered studies.

Second, it might be objected that our relatively small samples of effect sizes might lead either to Type II errors on the one hand or to making too much of statistically significant effects on the other. With regard to the possibility of Type II errors, we would note, for example, that even though a set of four pure RI design studies did not produce a significant moderating effect on proportion correct, a set of four mugshot commitment design studies did so, an effect that was judged to be trustworthy. Thus, small sample sizes, per se, in the present instance did not prevent relatively powerful moderator variables from emerging from our analyses. To guard against making too much of a statistically reliable effect for a moderator variable, we have proposed a reasonably stringent set of minimum criteria for what we have chosen to call a trustworthy effect. In addition, given the small sets of effect sizes that would be involved in any moderator analysis, we have imposed a limit of two moderator analyses for each dependent variable subjected to an overall

metaanalysis. In each instance, a second moderator analysis was conducted only after a careful analysis for the possibility of a confounding of the two variables yielded no discernable confounding. This latter step we have taken should have prevented the sort of overinterpretation of results that is based on multicollinearity.

## CONCLUSIONS

The results of our three metaanalyses and associated moderator analyses permit us to draw several conclusions concerning the impact of mugshot exposure on the fidelity of eyewitness memory. First, prior mugshot exposure decreases accuracy at a subsequent lineup, both in terms of reductions in rates for hits and correct rejections as well as in terms of increases in the rate for false alarms. Second, we have identified two variables, which are statistically reliable moderators of the effect of mugshot exposure, type of exposure design and number of mugshots exposed to the witness, the former being the more important of the two. At least with type of exposure design, there is some possibility of the criminal justice system making some beneficial modifications. Though police investigators do not perform experiments in the scientific sense, they could, for example, reduce the probability of untoward transference effects due to mugshot exposure by the simple expedient of not allowing a person to appear in a lineup whose photograph had been exposed in a prior set of mugshots, but not identified by the witness. Third, we have shown that findings from studies done to try to establish the reliability of the unconscious transference phenomenon and those from studies done to try to establish the reliability of the photobased lineup are both rightly considered as products of the same basic transference design. In both instances, there is transference of the identity of the target person to an innocent person, a person whose face occurred in proximate temporal context with that of the target. For transference studies included in our metaanalyses, when there were significant increases in transference errors, they were always associated with a proximate temporal context ranging from having the innocent person appear approximately 1 min before the target person to as long as a week afterward.

Clearly there would be benefits both in theory and application to further research efforts being aimed at exploring the boundaries of this zone of temporal proximity to the target person. Future research could also be profitably directed toward determining more precise parameters for the moderating effect of number of mugshots exposed to the witness.

## REFERENCES

*References marked with an asterisk indicate studies included in the metaanalyses.*

Bartlett, J. C., & Fulton, A. (1991). Familiarity and recognition of faces: The factor of age. *Memory & Cognition*, *19*, 229–238.

\*Brigham, J. C., & Cairns, D. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology*, *18*, 1393–1410.

\*Brown, E. L., Deffenbacher, K. A., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology*, *62*, 311–318.

- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior, 29*, 575–604.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- \*Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11*, 233–258.
- \*Cutler, B. L., Penrod, S. D., & O'Rourke, T. E. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour, 1*, 113–134.
- Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993).
- \*Davies, G., Shepherd, J., & Ellis, H. (1979). Effects of interpolated mugshot exposure on accuracy of eyewitness identification. *Journal of Applied Psychology, 64*, 232–237.
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687–706.
- \*Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 299–305.
- \*Dysart, J. E., Lindsay, R. C. L., Hammond, R., & Dupuis, P. (2001). Mugshot exposure prior to lineup identification: Interference, transference, and commitment effects. *Journal of Applied Psychology, 86*, 1280–1284.
- \*Gorenstein, G. W., & Ellsworth, P. C. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology, 65*, 616–622.
- Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326–338.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.
- Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the “general acceptance” of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56*, 405–416.
- \*Lindsay, R. C. L., Nosworthy, G. J., Martin, R., & Martynuck, C. (1994). Using mug shots to find suspects. *Journal of Applied Psychology, 79*, 121–130.
- \*Loftus, E. F. (1976). Unconscious transference in eyewitness identification. *Law and Psychology Review, 2*, 93–98.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252–271.
- \*Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition, 30*, 1219–1227.
- \*Perfect, T. J., & Harris, L. J. (2003). Adult age differences in unconscious transference: Source confusion or identity blending? *Memory & Cognition, 31*, 570–580.
- \*Read, J. D., Tollestrup, P., Hammersley, R., McFadden, E., & Christensen, A. (1990). The unconscious transference effect: Are innocent bystanders ever misidentified? *Applied Cognitive Psychology, 4*, 3–31.
- Reisberg, D., & Heuer, F. (2006). The influence of emotion on memory in forensic settings. In M. P. Toglia, J. D. Read, D. F. Ross, & Lindsay, R. C. L. (Eds.), *Handbook of eyewitness psychology: Memory for events* (Vol. 1). Mahwah, NJ: Erlbaum.
- Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043–1056.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192.
- \*Ross, D. F., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and mistaken identity: When a witness misidentifies a familiar but innocent person. *Journal of Applied Psychology, 79*, 918–930.
- Shapiro, P. N., & Penrod, S. D. (1986). A meta-analysis of facial identification studies. *Psychological Bulletin, 100*, 139–156.
- Simmons v. United States*, 390 U.S. 377 (1968).
- Stebly, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16*, 413–424.
- Williams, G. (1963). *The proof of guilt*. Springfield, IL: Charles C Thomas.
- Wise, R. A., & Safer, M. A. (2004). What U.S. judges know and believe about eyewitness testimony. *Applied Cognitive Psychology, 18*, 427–443.