

Witnessing-condition Heterogeneity and Witnesses' Versus Investigators' Confidence in the Accuracy of Witnesses' Identification Decisions

D. Stephen Lindsay, Elizabeth Nilsen, and J. Don Read

University of Victoria

Undergraduate participants were tested in 144 pairs, with one of each randomly assigned to a "witness" role and the other to an "investigator" role. Each witness viewed a target person on video under good or poor witnessing conditions and was then interviewed by an investigator, who administered a photo lineup and rated his/her confidence in the witness. Witnesses also (separately) rated their own confidence. Investigators discriminated between accurate and inaccurate witnesses but did so less well than witnesses' own confidence ratings and were biased toward accepting witnesses' decisions. Moreover, investigators' confidence made no unique contribution to the prediction of witnesses' accuracy. Witnesses' confidence and accuracy were affected in the same direction by witnessing conditions, and there was a substantial confidence-accuracy correlation when data were collapsed across witnessing conditions. Confidence can be strongly indicative of accuracy when witnessing conditions vary widely, and witnesses' confidence may be a better indicator than investigators'.

The degree of confidence displayed by witnesses giving identification testimony affects jurors' assessments of the credibility of the identification evidence (e.g., Cutler, Penrod, & Stuve, 1988; Leippe, Manion, & Romanczyk, 1992). This is an important finding, but only some of the witnesses who make suspect-identification decisions in police investigations ever testify in court. For example, witnesses who do not make a positive identification or who identify a known foil are unlikely to testify to their identification evidence. Even when a witness identifies a suspect, investigations are often abandoned for various reasons before they reach trial. Thus jurors see only some of the witnesses who take identification tests in forensic investigations, but the investigating officers see them all.

It is likely that police officers form opinions regarding the accuracy of witnesses' identification decisions, and such impressions may influence subsequent investigative efforts (e.g., encouraging or discouraging further investigation of a particular suspect). One aim of the current research was to explore participant/investigators' assessments of witnesses' accuracy. Another aim was to replicate and extend research by Lindsay, Read, and Sharma (1998), who argued that the relationship between

This research was supported by a Natural Sciences and Engineering Research Council of Canada grant to the first author and by a Natural Sciences and Engineering Research Council of Canada grant to the third author.

We thank Michael A. Hunter and Richard Lowry for consultations on statistical analyses of these data.

Address correspondence regarding this article to D. Stephen Lindsay, Department of Psychology, University of Victoria, P. O. Box 3050 STN CSC, Victoria, B.C., V8W 3P5, Canada. Email slindsay@uvic.ca.

witnesses' confidence and accuracy may be stronger across real-world cases than in most research. American Psychological Association, 12000. This paper is not the copy of record and may not exactly match the author's final manuscript. The final article is available, upon publication, at: <https://doi.org/doi/10.1023/A:1005504320565>

Why do lay people expect witnesses' confidence to be indicative of their accuracy? Lindsay et al. (1998) speculated that this expectation arises from everyday experiences demonstrating that individuals with expertise in a given domain tend to be both confident and accurate when tested within that domain, whereas those without expertise tend to lack both confidence and accuracy. If expertise is held constant, or varies only from almost none to scant, there is little intuitive reason to expect a strong confidence-accuracy (CA) relationship.

In most studies of the CA relationship, the CA correlation has been calculated either (a) for groups of participants who witnessed the target and were tested under a given set of conditions or (b) across conditions but with none of those conditions leading to high accuracy rates (e.g., see meta-analysis by Bothwell, Deffenbacher, & Brigham, 1987). When witnessing and test conditions are held constant only individual differences (in attention, motivation, history, etc.) and unintended procedural variations can lead to differences in witnesses' ability to identify the target. Consequently, in a given condition participants who respond correctly on the lineup test are likely to have only slightly better ability to identify the target than do participants who respond incorrectly, leading to a weak CA correlation. Even when the CA correlation is calculated across conditions, a strong relationship would be expected only if the manipulations endowed some witnesses with quite good ability to identify the target. For example, Cutler, Penrod, and Martens (1987) compared conditions ranging in accuracy from 17% (about chance) to 55%, and found that the CA collapsed across conditions was only $r = .20$. When nearly half of the participants in the best condition err, even those who respond correctly probably do not have particularly good ability. We propose that strong CA relationships are likely to emerge if some witnesses have quite good ability to identify the target and others do not.

In real-world cases, witnesses likely differ dramatically one from another in ability to identify the culprit, due to large differences across cases in witnessing and test conditions.¹ At one extreme, witnesses who had only a brief glimpse of a nondescript culprit under poor viewing conditions perform a difficult identification test months later. At the opposite extreme, witnesses who spent several minutes in well-lit, face-to-face interaction with a distinctive perpetrator are given an easy identification test within hours. Questions about the accuracy of witnesses' identifications arise across a wide range along this continuum (for examples of real-world cases involving challenges to identifications by witnesses who viewed the culprit under good conditions, see Connors, Lundregan, Miller, & McEwan, 1996; Shepherd, Ellis, & Davies, 1982). To the extent that witnessing and/or test conditions that greatly increase ability to identify the target also increase confidence, the CA relationship will be fairly strong across such cases.

Consistent with these ideas, Lindsay et al. (1998) found that manipulations of witnessing conditions that dramatically affected identification accuracy also affected confidence in the same direction. On average, the CA correlation for identifications within each of the four conditions used by

¹As pointed out by an anonymous reviewer of an earlier version of this manuscript, there is need for systematic research on how frequently identification evidence is collected in various kinds of cases and on how often (and under what circumstances) questions regarding identification accuracy arise.

Lindsay et al. was $r = .36$ (i.e., confidence accounted for only about 13% of the variation in accuracy within each witnessing condition). When witnessing-condition heterogeneity was increased by collapsing data across conditions, however, a substantial CA correlation emerged ($r = .59$), such that confidence accounted for about 35% of the variation in accuracy.

When conditions differ on a measure, variance on that measure will tend to be greater in data combined across conditions than in the individual conditions. This does not, however, mean that the finding of a stronger CA correlation across conditions than within conditions is a statistical artifact, because the correlation between two measures does not necessarily increase when variability in one of them increases. Suppose, for example, that the correlation between height and blue eye color is .20 both for men and for women, and that the prevalence of blue eyes is the same in the two sexes; under those assumptions, combining a sample of men with a sample of women would tend to increase variance in height but it would not tend to increase the correlation between blue-eyedness and height. If a variable (such as witnessing condition) affects two measures (such as confidence and accuracy) in the same direction, then the correlation between those measures will tend to be greater in data collapsed across those conditions than in the individual conditions, but this is not an uninteresting statistical artifact but rather the central point of the witnessing-condition heterogeneity argument: Real-world witnesses differ widely one from another in witnessing conditions, and witnessing conditions may often affect confidence and accuracy in the same direction, such that highly confident witnesses are substantially more likely to be accurate than non-confident witnesses.

Even with the data collapsed across witnessing conditions, the CA correlation reported by [Lindsay et al. \(1998\)](#) was far short of a perfect relationship. Moreover, a wealth of prior research demonstrates that manipulations can affect accuracy without affecting confidence and vice versa, which shows that confidence is not based solely on ability. Thus our claim is not that confidence is a perfect indicator when witnesses vary widely in ability to identify the target, but merely that it can be a substantially better indicator than most previous research implies.

Investigators' Assessments of Witnesses' Accuracy

As noted above, police investigators encounter a highly heterogeneous range of witnesses, from those who had only a fleeting glimpse of a nondescript perpetrator to those who spent substantial time in face-to-face interaction with a distinctive culprit. The [Lindsay et al. \(1998\)](#) findings suggest that such witnesses' self-rated confidence can be substantially predictive of their accuracy, at least if the identification test is conducted in a non-biased manner shortly after the witnessing event. To the best of our knowledge, police rarely ask witnesses for confidence ratings, but it seems likely that they often do make subjective assessments of witnesses' accuracy, and that such assessments can affect subsequent investigative efforts. This raises the question of how accurately investigators assess witnesses' accuracy.

Several studies have assessed participant-jurors' perceptions of witnesses' identification testimony (e.g., by showing them a videotape of a witness being cross examined about identification evidence) (e.g., [Leippe et al., 1992](#); [R. C. L. Lindsay, Wells, & Rumpel, 1981](#); [Wells & Leippe, 1981](#); [Wells, Lindsay, & Ferguson, 1979](#)). Such studies have generally found that participant-jurors display little ability to discriminate between accurate and inaccurate witnesses, tend to place undue faith in witnesses who viewed the target under poor conditions, and are overly affected by witnesses' apparent confidence. There are reasons to speculate, however, that police investigators have better insight than jurors into witnesses' accuracy: Investigators actively interact with witnesses (rather than passively observing their testimony) and they do so while the witness is making the initial identification

judgment (rather than long after the witness has committed to that judgment and been prepared for trial). Also, the witness may make any of a variety of responses during an initial identification test, whereas only witnesses who made a positive identification with at least a modicum of confidence are likely to testify in court or agree to be “cross examined” in a study.

The Current Experiment

Participants were tested in pairs, with one member of each pair randomly assigned to a “witness” role and the other to an “investigator” role. Each witness viewed a target person on video under good or poor witnessing conditions and was then interviewed by the investigator. The investigator administered a photo lineup identification test (which either included or did not include the target) and questioned the witness preparatory to rating the witness’s accuracy. Each witness also (separately) rated his/her own confidence. The procedure met all four of the recommendations proposed by Wells, Small, Penrod, Malpass, Fulero, and Brimacombe (1998): Foils were selected on the basis of fitting a description of the target, the investigator was blind to the identity of the target, witnesses were informed that the target might not be present in the lineup, and witnesses’ confidence was assessed immediately after the identification. We predicted that, as in [Lindsay et al. \(1998\)](#), witnessing condition would affect witnesses’ accuracy and confidence in the same direction, and that the within-condition CA correlations would be modest but that collapsing across conditions would reveal a substantial CA relationship. We also expected that investigators would discriminate between accurate and inaccurate witnesses, but that they would do so less well than witnesses’ own confidence ratings. Finally, we predicted that investigators would be biased toward accepting witnesses’ identification decisions as accurate.

Method

Participants

The participants were 288 University of Victoria undergraduate students who received a \$5 payment or optional extra credit in an introductory psychology course.

Materials

Each witness viewed one of two videos, each of which showed the same target (a young white man). One video was 10 s long and showed the target from one perspective, in one setting, wearing one outfit. The other video was 3 min long and depicted the target from a variety of perspectives, in several locations, and wearing several different outfits.

Witnesses were tested on one of four photo lineups, each consisting of eight head and shoulder frontal color photos of young white men wearing white lab coats and a “not present” sign. There were two target-absent and two target-present lineups: The two versions of each lineup type differed only in the location of the individual pictures. Foils closely fit a description of the target.

The rating scale for witnesses asked, “How confident do you feel that the person you picked in the photo lineup was the person you saw in the video?” (or, for those who chose the “not present” response alternative, how confident they were of the accuracy of that decision). The rating ranged from 1 (“purely guessing”) to 11 (“100% confident”). Investigators were asked, “How accurate do you feel the witness is?,” responding on a scale from 1 (“certainly inaccurate”) to 11 (“certainly accurate”).

Procedure

Participants were tested in pairs, with one member of each pair randomly assigned to the

witness role and the other to the investigator role. The investigator was given a clipboard with a photo lineup hidden beneath an instruction sheet, and sat outside the testing room and read the instructions, which explained that the witness was viewing a person on video and that subsequently the investigator was to administer a photo lineup identification test and question the witness with the aim of rating the witness's accuracy. Meanwhile, the witness was assigned to either the good or poor witnessing condition by a predetermined schedule. Witnesses in the good condition viewed the 3-min video, without any particular instructions, immediately after which the investigator entered the room and administered the lineup test. Witnesses in the poor condition were instructed to attend to the location in which the video was filmed; after viewing the 10-s video they completed a 2-min filler task and then wrote a brief description of the location in which the video was filmed before receiving the lineup test.

The investigator was outside the testing room for approximately 3 min in both the good and poor conditions, and was not informed of the witnessing-condition manipulation. The investigator entered the testing room and administered the lineup test by reading the instructions aloud and placing the lineup on a table in front of the witness. The instructions read by the investigator informed the witness that the man seen in the video might or might not be present in the lineup, and asked him/her to point to that man's picture or the "not present" sign. The investigator then asked the witness questions about the witness's confidence in the decision. Two sample questions were supplied on the investigator's instruction sheet ("How do you feel about your decision?" and "How easy was it for you to make your judgment?") and the instructions encouraged the investigator to ask any other questions that s/he thought might help assess the witness's accuracy.² The experimenter then provided the witness and investigator with clipboards with the printed confidence scales. Confidence ratings were made separately by the witness and investigator, who held up their clipboards so that they could not see each other's rating.

Results

We first report analyses of data from all 144 witness/investigator pairs, then describe findings for pairs in which the witness made a positive identification (as opposed to responding "not present"). For both sets of analyses, data were collapsed across the target-absent/target-present manipulation on the ground that in real-world cases investigators do not know whether or not the culprit is in the lineup (Wells, 1993).

All Witness Decisions

Witnesses' confidence and accuracy. Table 1 shows distributions of witnesses' confidence as a function of condition and accuracy. Responses were significantly more often accurate in the good witnessing condition (85%) than in the poor witnessing condition (33%), $\chi^2(1, N = 144) = 39.31, p < .01$. Also, confidence was significantly higher in the good witnessing condition ($M = 7.64, sd = 2.5$) than in the poor witnessing condition ($M = 4.08, sd = 1.95$), $F(1, 142) = 90.75, MSe = 5.02, p < .001$.

²Investigators' behaviors were not recorded. The experimenter's (EN) retrospective impression is that most investigators confined themselves to the sample questions or to slight paraphrases of them, but that some also asked witnesses a question along the lines of, "Why did you pick that person?"

Table 1: *Frequency Distributions of Witnesses' Confidence Ratings in all Witness Decisions.*

Condition	Confidence Rating										
	1	2	3	4	5	6	7	8	9	10	11
Good											
Accurate	0	0	3	1	7	0	12	8	13	7	10
Inaccurate	0	1	3	2	1	0	1	1	2	0	0
Poor											
Accurate	0	1	9	2	4	2	5	0	0	0	1
Inaccurate	1	12	13	9	7	2	1	2	1	0	0
Both											
Accurate	0	1	12	3	11	2	17	8	13	7	11
Inaccurate	1	13	16	11	8	2	2	3	3	0	0

Due to the skewed distributions, Spearman rank order correlation coefficients were used to assess the CA relationship (the same pattern of results was obtained when the data were analyzed with Pearson biserial correlation coefficients). Confidence and accuracy were significantly correlated in both conditions, and although the CA relationship was directionally stronger in the good witnessing condition ($r(70) = .37, p < .01$) than in the poor witnessing condition ($r(70) = .26, p < .02$), a Fisher r -to- z transform for the difference between two independent correlations indicated that this difference was not statistically reliable, $z = 0.72, ns$. Most important, when data were collapsed across conditions the CA correlation was a fairly substantial $r(142) = .55, p < .001$. The 95% confidence intervals for ρ , estimated with a Fisher r -to- z transform, were .43 to .65.

Investigators' confidence and witnesses' accuracy. Table 2 shows the distribution of investigators' confidence ratings as a function of condition and accuracy. Investigators' confidence was significantly greater in the good witnessing condition ($M = 7.94, sd = 2.64$) than in the poor witnessing condition ($M = 6.00, sd = 2.43$), $F(1, 142) = 21.00, MSe = 6.44, p < .001$. Investigators' confidence was reliably related to witnesses' accuracy in the good witnessing condition ($r(70) = .32, p < .01$) but not in the poor witnessing condition ($r(70) = .13, ns$). These values did not reliably differ, $z = 1.18, ns$. With data collapsed across conditions, the correlation between investigators' confidence and witnesses' accuracy was significant ($r(142) = .38, p < .001$; 95% confidence intervals = .23 to .51).

To test investigators' bias in assessing witnesses' decisions, investigators' confidence ratings of 1 (certainly inaccurate) to 5 (maybe inaccurate) were coded as "judged inaccurate," and those of 7 (maybe accurate) to 11 (certainly accurate) were coded as "judged accurate" (five pairs with ratings of 6 were deleted). Investigators were biased toward accepting witnesses' decisions as accurate: They correctly judged 73% of the accurate identification responses as accurate, but they correctly judged only 51% of the inaccurate responses as inaccurate. A test of the difference between two independent proportions showed that this bias was reliable, $z = 2.69, p < .01$.

Table 2: *Frequency Distributions of Investigators' Confidence Ratings in all Witness Decisions.*

Condition	Confidence Rating										
	1	2	3	4	5	6	7	8	9	10	11
Good											
Accurate	0	1	4	2	4	3	5	5	12	12	13
Inaccurate	0	0	2	1	3	0	2	1	1	1	0
Poor											
Accurate	0	0	5	2	4	0	5	0	4	2	2
Inaccurate	0	4	7	3	9	2	14	1	8	0	0
Both											
Accurate	0	1	9	4	8	3	10	5	16	14	15
Inaccurate	0	4	9	4	12	2	16	2	9	1	0

Witnesses' versus investigators' CA relationship. As can be seen by comparing the r values reported in the preceding two subsections, the correlation between investigators' confidence and witnesses' accuracy tended to be weaker than the corresponding correlation between witnesses' confidence and their own accuracy. This was the case directionally for all three comparisons, although the difference was statistically reliable only for the data collapsed across witnessing condition (the most forensically relevant as well as the most statistically powerful comparison), $t(141) = -3.13$, $p < .01$. For the data set as a whole, witnesses' confidence was twice as predictive of accuracy as was investigators' confidence (r^2 values of .30 and .15 for witnesses' and investigators' confidence, respectively).³ Investigators' and witnesses' confidence ratings were strongly correlated with each other (for the good and poor witnessing conditions, respectively, $r(70) = .60$ and $.69$, $ps < .001$; across conditions, $r(140) = .70$, $p < .001$).

Logistic regression analyses of witnessing condition, witnesses' confidence, and investigators' confidence as predictors of accuracy. The foregoing analyses reveal that witnessing condition, witnesses' confidence, and investigators' confidence were each related to witnesses' accuracy. They do not, however, indicate if these variables were predictive of accuracy independently of one another. To explore this question, we conducted a series of logistic regression analyses. Preliminary analyses indicated no significant 2- or 3-way interactions between witnessing condition, witnesses' confidence, and interviewers' confidence. A subsequent logistic regression with witnessing condition, witnesses' confidence, and interviewers' confidence indicated that the three predictors combined correctly

³ The confidence scale used by witnesses (from "purely guessing" to "100% confident") differed from that used by investigators (from "certainly inaccurate" to "certainly accurate"). The rationale for this difference is that it would not make sense for witnesses to be confident that they were wrong (if they were sure a response was wrong then they would make a different response), but investigators could sensibly believe that a witness was wrong. Nonetheless, the difference raises the possibility that the weaker CA correlation for the latter scale is an uninteresting measurement artifact. Against this possibility, the ranges of the two measures were virtually identical (see Tables 1-4) and the variance in confidence within each witnessing condition was directionally greater for investigators' confidence than for witnesses' confidence.

classified 75% of the subjects, $\chi^2(3, N = 144) = 57.942, p < .001$. That analysis revealed that witnessing condition ($\chi^2(1) = 8.68, p < .001$) and witnesses' confidence ($\chi^2(1) = 9.27, p < .001$) made unique contributions to predicting accuracy, whereas investigators' confidence did not ($\chi^2(1) = 0.94, ns$). The odds ratio for witnesses' confidence (1.463) indicates that the likelihood that a witness was correct increased nearly one and a half times for each 1-point increase along the 11-point witnesses' confidence scale.

Identifications Only

The analyses reported in this section must be interpreted with caution, due to the small number of accurate identifications in the poor witnessing condition (9 out of 43) and the small number of inaccurate identifications in the good witnessing condition (4 out of 33). These analyses are of particular interest, however, because identifications may have more serious real-world implications than responses of "not present."

Table 3: *Frequency Distributions of Witnesses' Confidence Ratings in Identifications.*

Condition	Confidence Rating										
	1	2	3	4	5	6	7	8	9	10	11
Good											
Accurate	0	0	1	0	3	0	5	4	5	4	7
Inaccurate	0	0	2	0	1	0	1	0	0	0	0
Poor											
Accurate	0	0	1	2	2	1	3	0	0	0	0
Inaccurate	0	7	13	6	2	2	1	2	1	0	0
Both											
Accurate	0	0	2	2	5	1	8	4	5	4	7
Inaccurate	0	7	15	6	3	2	2	2	1	0	0

Witnesses' confidence and accuracy for identifications. Table 3 shows the distribution of confidence ratings by witnesses who made identifications as a function of condition and accuracy. Identifications were significantly more often accurate in the good witnessing condition (88%) than in the poor witnessing condition (21%), $\chi^2(1, N = 76) = 33.474, p < .001$. Also as predicted, confidence in identifications was significantly higher in the good witnessing condition ($M = 8.03, sd = 2.51$) than in the poor witnessing condition ($M = 4.16, sd = 1.89$), $F(1, 74) = 58.91, MSe = 4.74, p < .001$. Confidence and accuracy were significantly correlated in both conditions, and although the CA relationship was directionally stronger in the good witnessing condition ($r(31) = .47, p < .01$) than in the poor witnessing condition ($r(41) = .32, p < .05$) this difference was not reliable, $z = 0.47, ns$. When the data were collapsed across conditions, the CA correlation for identifications was a very substantial $r(74) = .68, p < .001$ (95% confidence intervals = .54 to .79).

Investigators' confidence and witnesses' accuracy for identifications. Table 4 shows the distribution of investigators' confidence ratings for identifications as a function of condition and accuracy. Investigators were reliably more confident in witnesses in the good condition ($M = 8.30$, $sd = 2.30$) than in those in the poor condition ($M = 6.19$, $sd = 2.15$), $F(1, 74) = 17.04$, $MSe = 4.91$, $p < .001$. Investigators' confidence was reliably related to witnesses' accuracy in the good witnessing condition, $r(31) = .47$, $p < .01$, but not in the poor witnessing condition, $r(41) = .19$, *ns*. The difference between the two correlations was not statistically significant, $z = 1.32$, $p = .09$. Collapsing across witnessing conditions, the correlation between investigators' confidence and witnesses' accuracy was reliable, $r(74) = .50$, $p < .001$ (95% confidence intervals = .31 to .65).

As in the complete data set, investigators' bias was assessed by categorizing their confidence ratings as "judged inaccurate" (ratings of 1 to 5) or "judged accurate" (ratings of 7 to 11). Investigators had a substantial and significant bias toward accepting witnesses' identifications as accurate: 89% of accurate identifications were correctly perceived as accurate, but only 44% of false identifications were correctly judged inaccurate, $z = 3.93$, $p < .001$.

Table 4: *Frequency Distributions of Investigators' Confidence Ratings in Identifications.*

Condition	Confidence Rating										
	1	2	3	4	5	6	7	8	9	10	11
Good											
Accurate	0	0	1	0	1	3	3	3	3	10	5
Inaccurate	0	0	0	1	2	0	1	0	0	0	0
Poor											
Accurate	0	0	1	1	0	0	4	0	3	0	0
Inaccurate	0	3	4	0	6	2	13	1	5	0	0
Both											
Accurate	0	0	2	1	1	3	7	3	6	10	5
Inaccurate	0	3	4	1	8	2	14	1	5	0	0

Witnesses' versus investigators' confidence and accuracy of identifications. In the poor witnessing condition and for data collapsed across conditions, the relationship between witnesses' confidence and accuracy tended to be stronger than the relationship between investigators' confidence and witnesses' accuracy, although this difference was reliable only for data collapsed across condition, $t(73) = 2.79$, $p < .01$. Over both conditions, witnesses' confidence was nearly twice as predictive of the accuracy of their identifications as was investigators' confidence (r^2 values of .46 and .25 for witnesses' and investigators' confidence, respectively). Investigators' and witnesses' confidence ratings correlated with each other (for the good witnessing conditions, $r(31) = .56$, $p < .001$; for the poor witnessing condition, $r(41) = .71$, $p < .001$; across conditions, $r(72) = .72$, $p < .001$).

_____ *Logistic regression analyses of witnessing condition, witnesses' confidence, and investigators' confidence as predictors of accuracy of identifications.* These analyses paralleled those of the complete data set. There were no significant 2- or 3-way interactions between witnessing condition, witnesses' confidence, and interviewers' confidence as predictors of the accuracy of identifications. The three predictors combined yielded 75% prediction accuracy ($\chi^2(3) = 49.67, p < .001$). Witnesses' confidence ($\chi^2(1) = 5.52, p < .025$) and witnessing condition ($\chi^2(1) = 9.11, p < .001$) made significant unique contributions, whereas investigators' confidence did not ($\chi^2(1) = 0.26, ns$). The odds ratio indicated that the likelihood that a witnesses' identification was accurate increased 1.6 times for every unit increase in the confidence scale.

Discussion

Our findings regarding the relationship between witnesses' self-rated confidence and the accuracy of their identification decisions replicated the pattern reported by [Lindsay et al. \(1998\)](#). Specifically, when data from good and poor witnessing conditions were collapsed to create substantial heterogeneity across witnesses in ability to identify the target, witnesses' confidence was quite strongly correlated with their accuracy. The correlation was far from perfect—some non-confident witnesses were accurate and some confident witnesses were not—but the relationship was considerably stronger than it is generally said to be (e.g., for identifications, $r = .68$). These findings indicate that witnesses' self-rated confidence can have substantial value as an indicator of the accuracy of identification responses collected in a non-biased manner shortly after the witness viewed the to-be-identified person.

The substantial CA relationships obtained in this study and in some others (e.g., [Lindsay et al., 1998](#); [Read et al., 1997](#); [Sporer, Penrod, Read, & Cutler, 1995](#)) should not be construed as evidence that the confidence witnesses display in court predicts the accuracy of their identification testimony. Myriad social, cognitive, and statistical factors likely greatly attenuate the CA relationship over the months between when a crime was witnessed and when the witness testifies (see [Lüüs & Wells, 1994](#), and [Wells & Bradfield, 1998, 1999](#)). It is not to courtroom testimony to which the current findings may generalize, but rather to initial identification decisions made in non-biased testing situations shortly after a witnessed event.

The novel findings from this study are those concerning participant-investigators' confidence ratings. Investigators appeared to be sensitive to witnesses' confidence, in that witnessing condition had a reliable effect of investigators' confidence (even though investigators were not informed of the witnessing-condition manipulation) and investigators' confidence was correlated with witnesses' confidence. However, although investigators demonstrated some ability to differentiate between accurate and inaccurate witnesses, in general they did so more poorly than witnesses themselves. Indeed, the logistic regression analyses revealed that investigators' confidence made no independent contribution to the prediction of witnesses' accuracy. Investigators also revealed a bias toward accepting witnesses' decisions as accurate, and this bias was particularly evident for positive identification decisions.

It may be that trained police investigators working in real-world situations are better able to distinguish between accurate and inaccurate witnesses, and less biased toward accepting witnesses' judgments as accurate, than our participant-investigators. Formal training, the richness of naturalistic situations involving motivated witnesses, and extensive experience with numerous feedback opportunities (as when a witness confidently identifies a known foil) may enable forensic investigators to develop finely tuned perceptions of witnesses' likely accuracy. On the other hand, it may be that police tend to perform less well on these judgments than our participant-investigators did; our

participant-investigators explicitly questioned witnesses preparatory to making a quantitative rating of their confidence in the witness's response immediately after the identification judgment was given, whereas police may tend to assess witnesses' accuracy more informally, indirectly, and without making a record of that assessment. Moreover, our participant-investigators had no inside knowledge (e.g., no information about which if any member of the lineup was the "suspect") and no obvious motivation toward anything other than accuracy, whereas real-world police investigators may sometimes have considerable investment in the identification of a particular known suspect in the lineup.

In their influential "white paper" proposing guidelines for identification procedures, Wells et al. (1998) recommended that "a clear statement should be taken from the eyewitness at the time of the identification . . . as to his or her confidence" (p. 635). The rationale for this proposal was that although the CA correlation is weak it tends to be greater than zero for initial identifications, whereas post-identification processes such as social influence and commitment effects are likely to weaken the CA correlation over time: Given that witnesses' confidence is likely to affect jurors, Wells et al. argued, it is better that confidence be assessed immediately and formally at the initial identification rather than only much later and informally in court.

The current findings provide new evidence of the importance of assessing witnesses' confidence at the time of the identification, by showing that the CA relationship across highly heterogeneous witnesses can be quite strong at initial identification and that witnesses' self-rated confidence following identification decisions can be more diagnostic of accuracy than investigators' assessments. If the current findings are shown to generalize to more naturalistic situations, with realistic simulated crimes viewed by naive witnesses (and, ideally, with police experts conducting the identification interviews), the implications would be clear: Forensic investigators, blind to the identity of the suspect, should ask witnesses to provide quantitative confidence ratings immediately after making their identification judgments.

References

- Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72*, 691-695.
- Connors, E., Lundregan, T., Miller, N., & McEwan, T. (1996). Convicted by juries, exonerated by science: Case studies in the use of DNA evidence to establish innocence after trial. U.S. Department of Justice, Office of Justice Programs. <http://www.ncjrs.org/txtfiles/dnaevd.txt>.
- 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., & Stuve, T. E. (1988). Jury decision making in eyewitness identification cases. *Law and Human Behavior*, *12*, 41-56.
- Leippe, M. R., Manion, A. P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology*, *63*, 181-107.
- Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9*, 215-218.
- Lindsay, R. C. L., Wells, G. L., & Rumpel, C. M. (1981). Can people detect eyewitness identification accuracy within and across situations? *Journal of Applied Psychology*, *66*, 79-89.
- Lüüs, C. A., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and

- perseverance effects. Journal of Applied Psychology, *79*, 714-724.
- Read, J. D., Lindsay, D. S., & Nicholls, T. (1997). The relationship between accuracy and confidence in eyewitness identification studies: Is the conclusion changing? In C. P. Thompson, D. Herrmann, J. D. Read, D. Bruce, D. Payne, & M. P. Toglia (Eds.), Eyewitness memory: Theoretical and applied perspectives (pp. 107-130). Mahwah, NJ: Erlbaum.
- Shepherd, J. W., Ellis, H. D., & Davies, G. D. (1982). Identification evidence: A psychological evaluation. Aberdeen, Scotland: Aberdeen University Press.
- Sporer, S. L., Penrod, S. L., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. Psychological Bulletin, *118*, 315-327.
- Wells, G. L. (1993). What do we know about eyewitness identifications? American Psychologist, *48*, 553-571.
- Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect:" Feedback to eyewitnesses distorts their reports of the witnessing experience. Journal of Applied Psychology, *83*, 360-376.
- Wells, G. L., & Bradfield, A. L. (1999). Distortions in eyewitnesses' recollections: Can the postidentification-feedback effect be moderated? Psychological Science, *10*, 138-144.
- Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. Journal of Applied Psychology, *66*, 682-687.
- Wells, G. L., Lindsay, R. C. L., & Ferguson, T. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. Journal of Applied Psychology, *64*, 440-448.
- Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. Law and Human Behavior, *22*, 603-647.