
The "General Acceptance" of Psychological Research on Eyewitness Testimony

A Survey of the Experts

Saul M. Kassin
Phoebe C. Ellsworth
Vicki L. Smith

Williams College
University of Michigan
Northwestern University

ABSTRACT: Sixty-three experts on eyewitness testimony were surveyed about their courtroom experiences and opinions on various issues. There was a strong consensus indicated by an agreement rate of at least 80% that the data on the following topics are reliable enough to present in court: the wording of questions, lineup instructions, misleading postevent information, the accuracy-confidence correlation, attitudes and expectations, exposure time, unconscious transference, showups, and the forgetting curve. Over 70% of the experts also endorsed lineup fairness, the cross-race identification bias among White witnesses, and the tendency to overestimate the duration of events. Although most eyewitness experts who have testified have done so on behalf of criminal defendants, they were just as likely to consent for the prosecution as for the defense; moreover, they were more likely to agree to testify in civil cases than in criminal. Concerning their role in court, most respondents indicated that their main objective is to educate the jury, and that juries are more competent with the aid of experts than without. The results are discussed in relation to the "general acceptance" provision of the Frye test and the limitations of this test for determining the admissibility of expert testimony.

A few years ago, one of us was called by the defense lawyer in a burglary case and was asked to testify as an expert on eyewitness identification. After describing the cross-race identification bias (Brigham & Malpass, 1985), the effects of misleading postevent information (Loftus, Miller, & Burns, 1978), and the tenuous correlation between eyewitness accuracy and confidence (Wells & Murray, 1984), he was confronted on cross examination with quotes from isolated studies that had failed to replicate each of these phenomena. The prosecuting attorney concluded with the question: "Isn't it true, then, that there is always disagreement in your profession as to this area because it's in some ways still imprecise?"

Most psychologists who have taken the witness stand as experts have probably heard some version of this question. In fact, many have had their testimony excluded by a judge who decided that the research has not achieved an adequate level of scientific acceptance. Most judges (like most lawyers and jurors) lack the training necessary

to evaluate the merit of expert testimony. Over the years, however, case law and evidentiary standards have evolved to help them to make these decisions, and the level of agreement within the profession has emerged as a central criterion of acceptability.

This criterion was first enunciated in the 1923 case of *Frye v. United States* in which the defense sought to introduce the results of a "systolic blood pressure deception test," a forerunner of the polygraph. In its opinion, the D.C. Court of Appeals held that the test was inadmissible and pronounced that such scientific evidence "must be sufficiently established to have gained general acceptance in the particular field in which it belongs" (*Frye v. United States*, 1923, p. 1014). This standard is commonly referred to as the *Frye* test.

In the first case to propose standards for admitting expert testimony on eyewitness evidence, *United States v. Amaral* (1973), the *Frye* test was included as one of the following four requirements: First, the witness must be a qualified expert. Second, the testimony must concern a proper subject; that is, it should provide information that is not commonly understood by the average person. Third, following *Frye*, the testimony should conform to "a generally accepted explanatory theory" (*U.S. v. Amaral*, 1973, p. 1153). Finally, as with all other evidence (Federal Rules of Evidence, 1984, Rule 403), the probative value of the expert's testimony should outweigh whatever prejudicial effect it might have on the jury.

With regard to the first *Amaral* criterion, the Federal Rules of Evidence (1984, Rule 702) define an expert rather loosely as someone with relevant "knowledge, skill, experience, training, or education." In practice, the expertise of the witness is a matter of judgment over which trial judges have broad discretion. The second and fourth requirements, however, raise strictly empirical questions concerning the usefulness and impact of an expert's testimony on the jury.

In this latter regard, research on people's intuitions about eyewitness evidence indicates that although some factors are familiar to the average person, many are not. Studies using questionnaires (Brigham & Bothwell, 1983; Deffenbacher & Loftus, 1982; Yarmey & Jones, 1983) and mock trials (Cutler, Penrod, & Stuve, 1988; Lindsay, Wells, & Rumpel, 1981; for a review, see Wells, 1984)

indicate that common sense is not adequate preparation for the task. Moreover, research on the effects of expert testimony suggests that it may influence jurors in two ways. First, it leads them to scrutinize that evidence more carefully, to spend more time discussing it during deliberations, and to become more skeptical about its accuracy (Fox & Walters, 1986; Hosch, Beck, & McIntyre, 1980; Wells, Lindsay, & Tausignant, 1980). Second, it directs their attention to relevant aspects of the eyewitnessing situation as a basis for making their judgments (Cutler, Penrod, & Dexter, 1988). To date, it remains to be seen whether experts can enhance jurors' ability to distinguish accurate from inaccurate eyewitnesses, or whether the dangers of such testimony outweigh its probative value (e.g., whether jurors become not more or less skeptical, but more or less accurate in their judgments of eyewitness testimony).

The third *Amaral* requirement is that expert testimony be based on principles that are generally accepted within the relevant scientific community. Is there a consensus on the various propositions put forward by eyewitness experts? Trial and appellate courts disagree, sometimes concluding that there is (*United States v. Smith*, 1984) and sometimes concluding that there is not (*United States v. Fosher*, 1979). Before 1980, trial judges' decisions to exclude eyewitness experts nearly always were upheld. Since then, however, the research literature has grown at a rapid rate, and an increasing number of appellate courts have held that such experts should be allowed to testify.¹

Although it is not uncommon for legal scholars to look to judicial opinions in order to establish the general acceptance of scientific research (Giannelli, 1980, 1983), it seems more appropriate for us as psychologists to look to the scientific community itself. Unfortunately, there are scarcely any data on the degree of consensus existing among eyewitness psychologists. Some argue that many research findings are reliable enough to present in court (Buckhout, 1986; Loftus, 1983; Wells, 1986), and they regularly testify when asked. Others, equally vehement, argue that the research is inconclusive or inapplicable (Konecni & Ebbesen, 1986; McCloskey & Egeth, 1983; Pachella, 1986) and testify to that effect in order to obstruct the appearance of other experts. In the fireworks of the debate, it is important to keep in mind that the issue is not whether psychological research as a whole is reliable. Clearly, some propositions are more widely accepted than others (Bermant, 1986). Undiscriminating attacks on "psychological testimony" are no more illuminating than indiscriminating defenses and may be overly influential in a legal system that often takes a single dissenting voice as evidence for a lack of consensus (Ellsworth & Getman, 1987; Lempert, 1986). Like the reliability of the research findings themselves, their general acceptance is an empirical question.

In the only previous effort to address this question, Yarmey and Jones (1983) had 16 eyewitness experts read hypothetical situations designed to assess their opinions on a variety of topics and then asked them to choose one

out of four statements that best described the likely outcome. The results revealed high levels of agreement on many topics. Unfortunately, the perceived reliability of the findings was not addressed because the experts were not provided with a response alternative that enabled them to characterize the effects as weak, unreliable, or "hopelessly mixed" (Konecni & Ebbesen, 1986).

The present research was designed to determine what the experts think about various eyewitness phenomena. In doing so, we sought to update the Yarmey and Jones (1983) study, sample a larger number of eyewitness experts, address a broader range of issues, and improve on their method so that subjects would have a more exhaustive set of response options.

Method

Subjects

We compiled our list of experts by searching the eyewitness literature published in scientific journals between 1980 and mid-1986, including all senior authors and postdoctoral junior authors. This list was then supplemented by adding the authors of book chapters on eyewitness testimony, members of Division 41 of the American Psychological Association (APA) whose published work suggested an expertise in this area, and a few prominent psychologists whom we knew to be relevant experts on the basis of our own involvement in the field. This selection strategy resulted in a list of 119 prospective respondents, almost all psychologists affiliated with American, Canadian, and European institutions.

Sixty-three of the experts returned completed questionnaires for an overall response rate of 53%. Of the original sample of 119, however, six declined to participate on the grounds that they did not consider themselves experts.² Thus the response rate for those who were willing to characterize themselves as experts was 56%. Four experts—all conspicuous and fervent partisans in the con-

This research was conducted while Saul Kassin was on sabbatical from Williams College and was supported by a postdoctoral fellowship provided by the National Institute of Health to the Psychology-and-Law Program at Stanford University. Vicki Smith was supported by a graduate fellowship provided by the National Science Foundation. For part of the project, Phoebe Ellsworth was a Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, CA. She is grateful to the John D. and Catherine T. MacArthur Foundation, to the James McKeen Cattell Fund, and to the Center staff and Fellows for time, space, unsurpassable services, and intellectual stimulation. Last but not least, we are indebted to the 63 eyewitness experts who so generously took the time to complete our rather extensive questionnaire.

Correspondence concerning this article should be addressed to Saul M. Kassin, Department of Psychology, Bronfman Science Center, Williams College, Williamstown, MA 01267.

¹ Not all these cases address the general acceptance issue, however. In some, it is not reached (e.g., *State of Arizona v. Chapple*, 1983); in others, new standards are proposed as an alternative to the *Frye* test (e.g., *U.S. vs. Downing*, 1985).

² Some were reluctant because they had not kept up with the field; others because their own research was narrowly focused on a specific phenomenon and they were unfamiliar with the research on other topics.

Table 1
Eyewitness Topics and Statements Used to Describe Them

Topics	Statements
1. Stress	Very high levels of stress impair the accuracy of eyewitness testimony.
2. Weapon focus	The presence of a weapon impairs an eyewitness's ability to accurately identify the perpetrator's face.
3. Showups	The use of a one-person showup instead of a full lineup increases the risk of misidentification.
4. Lineup fairness	The more the members of a lineup resemble the suspect, the higher is the likelihood that identification of the suspect is accurate.
5. Lineup instructions	Police instructions can affect an eyewitness's willingness to make an identification and/or the likelihood that he or she will identify a particular person.
6. Exposure time	The less time an eyewitness has to observe an event, the less well he or she will remember it.
7. Forgetting curve	The rate of memory loss for an event is greatest right after the event, and then levels off over time.
8. Accuracy-confidence	An eyewitness's confidence is not a good predictor of his or her identification accuracy.
9. Cross-racial/white	White eyewitnesses are better at identifying other White people than they are at identifying Black people.
10. Cross-racial/Black	Black eyewitnesses are better at identifying other Black people than they are at identifying White people.
11. Postevent information	Eyewitness testimony about an event often reflects not only what they actually saw but information they obtained later on.
12. Color perception	Judgments of color made under monochromatic light (e.g., an orange streetlight) are highly unreliable.
13. Wording of questions	An eyewitness's testimony about an event can be affected by how the questions put to that witness are worded.
14. Unconscious transference	Eyewitnesses sometimes identify as a culprit someone they have seen in another situation or context.
15. Trained observers	Police officers and other trained observers are no more accurate as eyewitnesses than the average person.
16. Hypnotic retrieval	Hypnosis does <i>not</i> facilitate the retrieval of an eyewitness's memory.
17. Hypnotic suggestibility	Hypnosis increases suggestibility to leading and misleading questions.
18. Time estimation	Eyewitnesses tend to overestimate the duration of events.
19. Attitudes, expectations	An eyewitness's perception and memory for an event may be affected by his or her attitudes and expectations.
20. Sex differences	Women are better than men at recognizing faces.
21. Event violence	Eyewitnesses have more difficulty remembering violent than nonviolent events.

troversty over eyewitness experts—refused to participate on the grounds that the study was offensive or that “it depends” was the only honest answer they could give to the questions.³ These individuals were all men, and because they represented both poles of the dispute, we felt some reassurance that we had designed an unbiased questionnaire.

The Questionnaire

The questionnaire, which was anonymous (although many experts identified themselves), consisted of 24 pages and was divided into two parts. In the first part, we asked respondents to evaluate the reliability of 21 eyewitness phenomena. In the second part, we requested information about respondents' educational backgrounds, employment status, publication records, and courtroom expe-

riences. Accompanying the questionnaire was a brief, straightforward letter indicating the purposes of the study, as follows:

In view of the controversy surrounding the use of experts in eyewitness testimony, we are conducting a survey of what the experts actually think about various eyewitness phenomena. It is our hope that the results of this project will provide interested psychologists and the courts with a much-needed measure of the consensus in the scientific community. With that as our objective, we would like to include your views in this survey.

Specific opinion items. Based on recent reviews of the eyewitness literature (Lloyd-Bostock & Clifford, 1983; Shapiro & Penrod, 1986; Wells & Loftus, 1984), a list of 21 topics—some more established than others—was compiled. To minimize the risk that respondents would interpret the topics differently, each was specifically described in a sentence. Table 1 presents these topics and their descriptions.

For each topic, three kinds of questions were asked.

³ For each phenomenon included in our survey, the response options included the statement, “the research is inconclusive.”

First, respondents were asked how they would characterize the reliability of the phenomenon by checking one of the following seven alternative statements: (a) "The evidence suggests the reverse is probably true"; (b) "the evidence does not support it"; (c) "the evidence is inconclusive"; (d) "the evidence tends to favor it"; (e) "the evidence is generally reliable"; (f) "the evidence is very reliable"; and (g) "I don't know." Next we asked, "Would you say that most jurors believe the statement to be true as a matter of common sense? (yes, no, don't know)." Third, we asked whether the research evidence met the expert's standard of reliability for presentation in court. The specific questions were as follows: (a) "Do you think this phenomenon is reliable enough for psychologists to present it in courtroom testimony?" (b) "Under the right circumstances, would you be willing to testify in court about this phenomenon?" (c) "Have you ever actually testified as to your opinion of this phenomenon?"

Personal background information. Following the opinion questions, respondents were asked about their professional credentials (degrees obtained, primary area of specialty within psychology, membership in Division 41 of the APA, and current employment), their relevant scholarly achievements (number of eyewitness publications in scientific journals, law reviews, books, chapters, and newsletters), and their courtroom experience (the estimated number of times they were asked to testify, the number of times they agreed to testify, and the number of times they actually testified). With regard to the last inquiry, we asked whether they had been called on a criminal or civil case and whether they were contacted by the prosecution (plaintiff) or defense. Two general questions then were asked—both on topics of controversy among eyewitness experts. First, "What do you see as the primary role of the eyewitness expert—to educate the jury, assist a particular party, or other (please specify)?" Second, "In general, would you say that juries are better equipped to evaluate eyewitness testimony with or without the aid of a competent expert—with, without, or no difference?" Finally, respondents were asked to report on eyewitness topics on which they had testified that were

not covered in the questionnaire, and if they had ever refused to testify, to indicate the reasons for their refusal.

Results and Discussion

The Experts

Among our 63 respondents, 60 had PhDs in psychology (six of whom also had JDs), two were doctoral students, and one was an MD. Forty-eight of the respondents (76%) were members of Division 41 of the APA. Thirty-one (49%) identified themselves as social or personality psychologists; 17 (27%) were in perception and cognitive psychology; 9 (14%) were in clinical; and 5 (8%) were in "other" areas of the discipline. Regarding their scholarly credentials, 47 (75%) of our experts reported having authored or coauthored one or more publications on the subject of eyewitness testimony. Overall, respondents had a mean of 6.35 relevant publications each. Of these, 65% were scientific journal articles, 20% were book chapters, 9% were magazine and newsletter articles, 3% were law review papers, and 3% were books containing relevant material.

Turning to our experts' courtroom experience, we found that 34 (54%) had testified on the subject of eyewitness testimony at least once. Overall, they estimated having testified on 478 occasions ($M = 7.59$): 364 times for the criminal defense, 29 times for the criminal prosecution, 54 times for the civil plaintiff, and 31 times for the civil defense. Thus, when eyewitness experts appear in court, it is usually (76% of the time) on behalf of a criminal defendant. Does this lopsided involvement reflect a liberal bias within the profession, or does it merely reflect differential demands from within the legal community? To address this question, we compared the frequency with which the experts in our sample were asked to testify, had agreed, and had actually testified in court.

The results were instructive. Table 2 shows that experts were asked to testify an estimated 1,268 times and that 1,063 (84%) of these requests were made by criminal defense attorneys. Clearly, the major demand for eyewitness

Table 2
Estimated Number of Times Experts Were Asked, Had Agreed, and Actually Testified in Court

Action	Criminal				Civil				Totals	
	Prosecution		Defense		Plaintiff		Defense		N	%
	N	%	N	%	N	%	N	%		
Asked to testify	56	—	1,063	—	96	—	53	—	1,268	—
Agreed to testify	44	79 ^a	755	71 ^a	87	91 ^a	49	92 ^a	935	74 ^a
Actually testified	29	66 ^b	364	48 ^b	54	62 ^b	31	62 ^b	478	51 ^b
Total yield		52 ^c		34 ^c		56 ^c		58 ^c		38 ^c

^a These numbers represent the experts' rate of agreement (i.e., the percentage of times they agreed to testify upon request).

^b These numbers represent the percentage of experts who had agreed who ultimately testified in court.

^c These numbers represent the percentage of experts initially asked who ultimately testified.

ness experts comes from the criminal defense bar. However, the experts themselves showed no preference for one side or the other in criminal cases. They were just as likely to agree to appear for the prosecution as they were for the defense. If the experts showed a preference for a certain kind of involvement, it was for civil cases, rather than criminal (91.3% and 71.4% agreement rates, respectively, $\chi^2(1, N = 63) = 26.81, p < .001$).

Thirty-eight of our experts indicated that on at least one occasion they had refused a request for testimony. Most of their reasons fell into four categories. Sixteen experts had refused on personal or moral grounds. These decisions were based on mistrust of the lawyer; a fear of being used as a "pawn," or not being permitted to qualify answers; a belief that the evidence was overwhelming and the defendant was guilty; or an emotional distaste for the case itself (e.g., child sex abuse or a Nazi deportation case). Fifteen experts cited the belief that they had nothing relevant or useful to say on behalf of the party that called (e.g., the identification procedures seemed fair; the testimony would not benefit the jury). Seven experts cited a lack of preparation time either because of their own busy schedules or because they were contacted too late. Five experts cited their lack of expertise on a particular subject, as when a social psychologist was asked to testify on certain aspects of visual perception. The remaining reasons did not fit into these categories (i.e., the stress of cross-examination, or the party's unwillingness to pay the expert's consulting fee).

Judgments of Eyewitness Phenomena

Opinions of the specific eyewitness topics were analyzed with four goals in mind: (a) to conduct the *Frye* test by describing how the scientific community evaluates each phenomenon, (b) to explore the relation between the experts' assessment of the reliability of a finding and their willingness to testify on it, (c) to compare different kinds of experts on these issues, and (d) to measure experts' personal opinions about what phenomena jurors understand as a matter of common sense.

Table 3 presents the distribution of responses for experts' perceptions of each of the 21 empirical propositions. Perhaps more important from a forensic standpoint, Table 4 shows—item by item—the proportion of experts who believed the evidence reliable enough to be presented in court, were themselves willing to testify on the topic, already *had* testified on the topic, and believed that jurors understood it as a matter of common sense.

Because the courts have never offered an explicit or quantitative definition of general acceptance and because the notion of general acceptance as a steady state makes little scientific sense, it would be foolish to try to establish a list of eyewitness phenomena that "pass" or "fail" the *Frye* test. Nevertheless, both the absolute percentages and the rankings are informative. Table 4 shows that nine propositions were viewed by at least 80% of the experts as reliable enough to be presented in court. In contrast,

Table 3
Distribution of Reliability Judgments for the 21 Eyewitness Topics

Topic	Reverse	No support	Inconclusive	Tends to favor	Generally reliable	Very reliable	Don't know or missing
1. Stress	0	0	10	25	22	3	3
2. Weapon focus	0	0	17	25	15	3	3
3. Showup	0	0	5	9	13	28	8
4. Lineup fairness	6	6	3	10	14	15	9
5. Lineup instructions	0	0	1	10	22	30	0
6. Exposure time	0	0	4	14	18	22	5
7. Forgetting curve	2	1	5	10	9	29	7
8. Accuracy—confidence	0	1	6	9	18	28	1
9. Cross-racial—White	0	2	5	15	23	17	1
10. Cross-racial—Black	1	6	15	16	15	5	5
11. Postevent information	0	1	0	13	20	28	1
12. Color perception	0	0	2	3	8	12	38
13. Wording of questions	0	0	0	3	14	46	0
14. Unconscious transference	0	0	3	23	21	11	5
15. Trained observers	1	2	6	19	21	2	12
16. Hypnotic retrieval	4	4	21	9	13	3	9
17. Hypnotic suggestibility	0	0	12	14	11	17	9
18. Time estimation	0	0	4	8	21	16	14
19. Attitudes and expectations	0	0	1	18	16	27	1
20. Sex differences	0	16	14	11	4	0	18
21. Event violence	2	3	10	19	11	2	16

Table 4
Discrete Judgments and Opinions Concerning the 21 Eyewitness Topics

Topics	Reliable enough?	Would you testify?	Have you testified?	Common sense?
1. Wording of questions	96.8	88.9	20.6	26.8
2. Lineup instructions	95.1	85.5	25.8	36.7
3. Postevent information	87.1	82.5	41.3	7.5
4. Accuracy-confidence	87.1	82.5	36.5	3.2
5. Attitudes and expectations	86.9	73.0	19.0	40.8
6. Exposure time	84.7	72.1	26.2	89.1
7. Unconscious transference	84.5	67.2	29.0	17.0
8. Showups	83.1	77.0	23.0	28.6
9. Forgetting curve	82.5	78.3	27.9	23.9
10. Cross-racial/White	79.4	71.4	31.7	60.0
11. Lineup fairness	77.2	69.5	27.1	51.2
12. Time estimation	74.5	57.4	20.0	4.7
13. Stress	70.5	64.5	38.1	24.0
14. Hypnotic suggestibility	68.5	55.2	8.6	21.4
15. Color perception	65.7	32.7	9.6	32.0
16. Trained observers	58.7	51.7	20.0	3.5
17. Weapon focus	56.5	53.3	27.4	11.4
18. Hypnotic retrieval	51.9	46.4	8.8	2.0
19. Cross-racial/Black	48.3	43.3	20.0	50.0
20. Event violence	36.0	30.4	14.0	9.8
21. Sex differences	11.1	13.0	00.0	8.3

Note. Numbers represent the percentage of experts who responded affirmatively to each question. The topics are rank-ordered according to their scores on the question, "Do you think this phenomenon is reliable enough for psychologists to present it in courtroom testimony?"

fewer affirmative answers were obtained when respondents were asked if they themselves would testify, $\chi^2(1, N = 63) = 26.1, p < .001$. On this question, only four propositions elicited 80% willingness from the experts.

At the heart of the debate between those who favor and those who oppose the use of eyewitness experts is a dispute over how clear and convincing the research support must be before it is offered in court (McCloskey, Eggeth, & McKenna, 1986). In order to address this question, we associated each respondent's estimate of the reliability of each statement with that same respondent's judgment of whether it was reliable enough for testimony. By combining the data across all respondents and topics, we derived the distribution presented in Table 5. The results were rather straightforward. The experts in our sample seldom (17%) reported a willingness to testify on topics if they felt the data were "inconclusive." A larger percentage (32%) said they would testify when there was "no support," presumably to say just that. The experts were split about what to do when the evidence "tends to

favor" a topic or "suggests the reverse is probably true" (56% and 60%, respectively). Of those who described the evidence as "generally reliable," however, 80% said they would testify; of those who described it as "very reliable," 92% said the same.

How reliable is reliable enough? The results suggest two answers. First, the vast majority of experts said they were willing to testify only on research findings they considered at least generally reliable. Indeed a statement we had included that is completely without support (that women are better than men at recognizing faces; see Table 1, item 20) was most often rejected. Second, a small number of experts expressed a reluctance to testify even when they consider a finding to be reliable. This conclusion follows from the fact that on 20 of the 21 topics, fewer respondents said they would testify than evaluated the evidence as reliable enough for psychological testimony in general (see Table 4). This discrepancy indicates that, for at least some psychologists, the decision to testify is based not only on scientific reliability but also on other considerations, such as external validity (Konecni & Ebbesen, 1986; Pachella, 1986), an apprehension over the effects on the jury (McCloskey & Eggeth, 1983), or simply personal lack of interest in appearing in court.

On the question of common sense, respondents' opinions seem well informed by the relevant research. As shown in the far right-hand column of Table 4, for example, most of our experts believe that jurors are familiar with the effects of exposure time and unfamiliar with the inadequacy of confidence as a predictor of accuracy or the potentially biasing effects of misleading postevent information. Opinions were mixed on jurors' intuitions about cross-racial identification and lineup fairness. These results closely parallel studies of people's actual intuitions (Deffenbacher & Loftus, 1982; Hastie, 1980; Yarmey & Jones, 1983; for a review, see Wells, 1984). Also, the present data suggest that the experts disagree with McCloskey and Eggeth's (1983) argument that it is difficult to find eyewitness research that is both reliable and nonintuitive. Table 4 indicates that they identified the wording of ques-

Table 5
Percentage of Experts Who for Each Topic Judged It to Be Reliable Enough and Would Themselves Testify

Opinions on reliability	Judgments	
	Reliable enough?	Would you testify?
1. Reverse is probably true	60.0 _b	60.0 _c
2. No support	30.0 _c	31.7 _d
3. Inconclusive	8.6 _d	17.3 _e
4. Tend to favor	64.1 _b	55.8 _c
5. Generally reliable	93.6 _a	79.6 _b
6. Very reliable	99.4 _a	91.9 _a

Note. Under each column, percentages not sharing a common subscript differ at $p < .05$.

tions, lineup instructions, postevent information, the accuracy–confidence correlation, unconscious transference, showups, and the forgetting curve as factors that are widely accepted by the scientific community but relatively unknown to the general public.

Individual Differences

According to Rule 702 of the Federal Rules of Evidence, experts are loosely qualified “by knowledge, skill, experience, training, or education.” One might, therefore, expect to find a good deal of variability among psychologists who could be considered eyewitness experts. To identify possibly relevant subsamples, respondents were classified according to their credentials on two independent dimensions: research experience (46 had authored at least one eyewitness publication; 17 had not) and courtroom experience (34 had testified on at least one occasion; 29 had not). Of interest was whether the resulting subsamples differed in their judgments (i.e., the measures shown in Table 4).

For each respondent, percentage scores were derived by dividing their number of affirmative responses by the total number of responses they made to Questions 3 (Is it reliable enough?) and 4 (Would you testify?) of the questionnaire. Two-way ANOVAS (publication status \times witness status) on these scores revealed strong main effects for both categorical variables. As one might expect, respondents who had courtroom experience judged more propositions to be reliable enough for testimony ($M_s = 78.21$ compared to 65.37), $F(1, 59) = 4.76$, $p < .05$, and were willing to testify more often about these propositions ($M_s = 74.52$ compared to 48.43), $F(1, 59) = 14.68$, $p < .001$. The effects of publication status were somewhat more complicated: Those who had published judged fewer statements as reliable enough for testimony ($M_s = 69.84$ compared to 81.30), $F(1, 59) = 5.01$, $p < .01$, but were themselves *more* often willing to testify ($M_s = 64.93$ compared to 55.18), $F(1, 59) = 5.10$, $p < .01$. One possible interpretation of this pattern is that one’s willingness to testify depends not only on one’s faith in the research but also on one’s own ability to describe that research accurately. Although active researchers appear to set more stringent standards, perhaps they also are more certain of their own competence.⁴

Opinions on Eyewitness Experts and the Jury

Much recent debate has centered on the role assumed by expert witnesses. According to one position, because experts are almost always called by one side in a dispute, they should consider themselves advocates for that side. According to the alternative position, experts should instead maintain a more neutral posture as impartial ed-

ucators (see Loftus, 1986). We asked whether the primary role of the eyewitness expert was to educate the jury or assist a particular party (an “other—please specify” option was included). Out of 58 respondents who answered the question, 52 (90%) selected the educator role; only three (5%) selected the advocacy role; two “other” respondents indicated that their primary objective was to improve procedures used by the police for obtaining eyewitness accounts. Obviously, one cannot infer from these results how experts would actually behave once they were on the witness stand. If there is a single point on which Loftus (1986) and McCloskey et al. (1986) agreed, it was that the pressures of the adversarial situation often lead scientists to experience a good deal of role conflict. Still, it is interesting that the vast majority of experts aspire to the more difficult role of educator.

Finally, we asked respondents to consider the impact of expert testimony on jury decision making. Loftus (1983) and others have maintained that because jurors overbelieve eyewitness testimony, the presence of an expert has the desirable effect of leading them to scrutinize the evidence more carefully. On the other hand, McCloskey and Eggeth (1983) argued that experts do more harm than good, perhaps confusing juries, misleading them, and making them *too* skeptical of eyewitness evidence, without improving their ability to discriminate accurate from inaccurate witnesses. At this point, research has not provided definitive answers to these empirical questions (see Wells, 1986). Nevertheless, most of our respondents (89%) believed that juries are better off with the aid of a competent expert; 3% said juries are better without experts; and 8% thought it makes no difference.

General Discussion

The present survey revealed some interesting facts about eyewitness experts, their appraisals of the research literature, and their views on their own role in court. To begin with, there was no evidence to support the fear that eyewitness experts are inherently biased in favor of criminal defendants; indeed, they are just as likely to agree to testify on behalf of the criminal prosecution, and they are much more likely to become involved in civil cases. The disproportionate frequency with which the experts testify for the criminal defense merely reflects the fact that eyewitness testimony is most often introduced by the prosecution and countered by the defense.

Concerning the general acceptance provision of the *Frye* test, the survey yielded clear results. There was a strong consensus, indicated by an agreement rate of at least 80%, that the following topics are sufficiently established for psychologists to discuss in court: the wording of questions, lineup instructions, the effects of misleading postevent information, the accuracy–confidence correlation, attitudes and expectations, exposure time, unconscious transference, showups, and the forgetting curve. If a 75% agreement rate were deemed sufficient, then the cross-racial identification bias among White witnesses, lineup fairness, and the tendency to overestimate the duration of events would also be included on the list. On

⁴ It could be argued that only those psychologists with a record of publications should be counted as part of the relevant scientific community. The main results from Table 4 were thus derived for this more selective subsample ($N = 47$). By and large, however, the main findings (i.e., the rank ordering of the 21 phenomena) were unchanged. These results are available upon request.

establish "general acceptance." In practice, that evidence might consist of a haphazard list of scientific publications, prior judicial opinions, or the statements of one or two experts in the field (Black, 1988; Giannelli, 1980). Indeed, any attorney who wants to block the introduction of an expert can find a counter-expert to claim that the findings are not well enough established. The opinion of one person with a PhD, of course, tells us little about the larger community of experts.

By providing far better evidence of consensus, our survey should prove useful not only to judges but also to eyewitness experts and trial lawyers. As far as the experts themselves are concerned, the results should offer guidance as to the appropriate contents of their testimony. Cross examiners too could use these data to discredit experts who stray onto topics that are not generally accepted. In the long run, the information provided by this survey should lead to expert testimony that more accurately represents opinions within the scientific community at large. Indeed, comparable surveys in other areas of psychological testimony (e.g., rape trauma syndrome; see Frazier & Borgida, 1988) stand to make the same contribution.

The fact that we offer the best evidence on the general acceptance issue, however, does not mean that we approve of "general acceptance" as a criterion for the admissibility of expert testimony. In the first place, if the most important criterion for expert testimony is that it assist the trier of fact, then its content is not the only factor to be considered. Following common Bayesian logic, in which the impact of new information is measured against existing beliefs, jurors' prior beliefs must also be weighed in deciding whether the testimony will prove helpful. As such, there are four possible outcomes. To illustrate these possibilities, we draw on Hastie's (1980) study of the eyewitness factors actually discussed during mock jury deliberations.

The first possibility is that jurors are ignorant of phenomena that are widely accepted by the experts. Post-event information is a good example. The fact that new information can influence a witness's testimony is well documented (Wells & Loftus, 1984). Most of our experts agree on the reliability of this effect (87.1%), and yet very few believe that it is common knowledge among jurors (7.5%). It appears they are right. When Hastie coded 11 deliberations for all references to human memory, he found that "the possibility of distortion, elaboration or additions to a witness's memory trace" was mentioned in only one group. Thus, in the absence of an expert, jurors do not generally consider the effects of postevent information.

The second possibility is that there is consensus among the experts, but that jurors too are aware of the phenomenon. In other words, the research evidence on which the experts rely is a matter of common sense (McCloskey & Egeth, 1983; *United States v. Fosher*, 1979). For example, 84.7% of our experts believe that "the less time an eyewitness has to observe an event, the less well he or she will remember it," but 89.1% also believe that

jurors are by themselves aware of this effect. In Hastie's study, the duration of the crime events came up in 10 of the 11 juries that discussed eyewitness testimony. On this topic, then, expert testimony is not helpful to the jury because the jury needs no help.

The third possibility is that neither the experts nor the jurors have a basis for trusting a phenomenon because the empirical evidence is too weak or scanty to pass the *Frye* test. In our study, for example, most experts (64%) were unwilling to endorse the statement that "eyewitnesses have more difficulty remembering violent than nonviolent events." If jurors are also uncertain about the role of violence, then presumably an expert's testimony would either be useless (if he or she testified that the data are ambiguous) or misleading (if he or she claimed that violent events were difficult to remember).

The fourth possibility is that jurors believe a proposition, but the research results are ambiguous, contradictory, or nonexistent. There may be a consensus among the scientific community that common sense is not generally true, as with the proposition that confidence is a good predictor of accuracy (Wells & Murray, 1984). Or there may be disagreement among scientists, as there is on the effects of high stress on memory. Yet, Hastie found that the accuracy-confidence assumption and the assumption that stress facilitates memory (1980, p. 9) were both very common, each arising in 10 deliberations. In such cases in which jurors are convinced that a proposition is true, information concerning the *lack* of consensus could prove extremely helpful.

In short, if we accept that the basic purpose of expert testimony is to assist the trier of fact, then its usefulness depends not only on a general acceptance (or lack thereof) among scientists, but on people's prior beliefs as well.

There are other problems with using consensus as a criterion for admissibility. The rates of agreement in our survey were not, in fact, perfect measures of the research. For example, there are no studies that compare lineups and showups (Item 3), but a good deal of theory and research on the functional size of lineups leads one to expect that showups are likely to be suggestive. By agreeing that showups increase the risk of misidentification, the experts were thus extrapolating from related work. On the other hand, the effects of monochromatic light on color perception (Item 12) are very well-documented and familiar to anyone who studies color vision. Yet our experts exhibited relatively low levels of consensus on this item, and a large number of "I don't know" responses. Clearly, our sample of respondents—many of whom are specialists in social cognition, not color vision—was not the most appropriate possible sample for *this item*. This finding thus exposes yet another problem with the general acceptance standard: It may be as difficult to define the relevant scientific community as it is to define general acceptance itself.

Finally, our results should not be taken to imply that using psychological experts is the best possible solution for the problems arising from eyewitness testimony. Walker and Monahan (1987; see also Monahan & Walker,

the other hand, it is noteworthy that the effects of stress, hypnosis, weapon focus, trained observers, event violence, and the cross-racial bias among Black witnesses did *not* elicit high degrees of consensus; as expected, neither did the item on sex differences, which was completely without support but was included as a means of evaluating whether respondents were discriminating in their judgments.

Unlike previous research, our study provided the experts with ample opportunity to express a lack of confidence in the various propositions. They could reject the effects as unreliable by characterizing the results as "inconclusive" or "without support." We also differentiated between general scientific reliability and reliability for courtroom purposes, and we found high levels of consensus for both. Thus, it cannot be argued that our experts felt the results were reliable enough for the development of science but not sufficiently generalizable to be helpful to juries (Pachella, 1986). Fully 89% of our respondents believed that juries are more competent with assistance from an eyewitness expert than without. This finding cannot be easily dismissed as a matter of self-justification: Seventeen of the 19 respondents who had published research but had *never* testified shared this belief.

We also attempted to examine the standards of reliability eyewitness experts demand before agreeing to testify. Basically, the issue consists of two questions: (a) How much research evidence does it take for experts to conclude that a finding has support? (b) How much support must an expert perceive before he or she is willing to present that finding in court? Although we did not address the important first question, we did address the second. By examining the association between experts' perceptions of reliability and judgments of their own willingness to testify, we found that most experts say they would agree to testify on phenomena they consider "generally" or "very reliable." At the same time, 56% reported a willingness to testify when the research evidence merely "tends to favor" a finding, and 17% were willing when the evidence was "inconclusive." Unfortunately, our data do not tell us what these experts would actually say about findings they perceived as relatively weak.

Limitations of the Study

Our survey is limited in four ways. To begin with, not all possible eyewitness experts were included in the sample; 44% of those on our initial list did not return their questionnaires. Also, the original list undoubtedly omitted an unknown number of experts who have testified but have not published in the eyewitness area, or who have published in areas that were missed in our search of the literature.

Second, the results are based entirely on self-report data. For obtaining estimates of courtroom and publication experiences, the survey method is fine, and for assessing the consensus of opinion on various eyewitness findings, it is the only plausible method and is far better than other means of establishing general acceptance (Giannelli, 1980). One should be more cautious however,

when it comes to the question of behavioral intentions (i.e., *would* you testify?).

Third, our instrument did not include all possible factors that have been studied or presented by eyewitness experts. We tried to include topics frequently raised in court, and we tried to cover a range from well-documented to poorly documented findings. Inevitably, certain potentially important stimulus variables (e.g., disguise), storage variables (e.g., exposure to mugshots), and retrieval variables (e.g., context reinstatement) were omitted (see Cutler, Penrod, & Martens, 1987). In fact, in response to our inquiry concerning topics not covered in the questionnaire, several interesting omissions were noted by our experts.⁵ Finally, most of our items described main effects. Specific interactions—for example, that the magnitude of the accuracy–confidence correlation depends on the quality of the viewing conditions (Bothwell, Deffenbacher, & Brigham, 1987)—were not included (the one exception is the cross-racial bias, where the interaction between race of witness and culprit was separated into two statements).

Fourth, *any* attempt to assess the general acceptance of scientific knowledge is necessarily limited by the passage of time. As research accumulates, as new factors are discovered, and as the effects of already known factors are modified, expert opinion is likely to change. We appear to have stumbled onto a good example of a phenomenon that would probably have fared differently today than it did just two years ago. During their initial exchange, McCloskey and Egeth (1983) and Loftus (1983) disagreed over the reliability of the weapon focus effect, that the presence of a weapon draws a witness's attention and thus impairs recognition of the culprit's face. McCloskey and Egeth (1983) pointed out that there was virtually no solid evidence for the phenomenon. In response, Loftus (1983) argued that the effect can be inferred from more traditional studies of attention to novel objects. No doubt aware of the issue, the experts in our sample split about 50–50 on the reliability of this phenomenon. Yet, one wonders how these experts would respond today to the same issue. Over the past year, the weapon focus effect has received strong support in three independent investigations (Cutler, Penrod, & Martens, 1987; Loftus, Loftus, & Messo, 1987; Tooley, Brigham, Maass, & Bothwell, 1987). As this example illustrates, research on expert opinion should be periodically updated to account for developments in the state of our knowledge.

Limitations of the "General Acceptance" Standard

Until now, one of the problems with the *Frye* test—as applied not only to psychology but to other scientific evidence as well—has been that the courts offer little guidance on the kinds of evidence that might be used to es-

⁵ Six experts had testified on biased retrieval procedures (e.g., exposure to mugshots), four on "earwitness" testimony (e.g., voice identification, memory for conversations), four on visual perception (e.g., dark adaptation, distance and size estimation), three on state-of-mind issues (e.g., effects on memory of drugs, alcohol, and chronic life stress), three on the competence of children as witnesses, and five on miscellaneous topics.

1988) recently suggested that because expert testimony is costly, an alternative would be to educate juries through cautionary instructions. Such instructions should be based on research presented to the court in briefs or on the court's own review of the literature. Such an instruction has, in fact, been in use ever since it was adopted in the 1972 case of *United States v. Telfaire* (see Bibicoff, 1983). Although recent mock-jury research indicates that it does not have the desired impact (Cutler, Penrod, & Dexter, 1988) and that it needs to be written in more accessible language (Greene, 1988), it is unclear whether the problem is inherent in the use of jury instructions per se as a medium of communication or whether the specific content of the *Telfaire* instruction is flawed. With that problem in mind, the present study should assist the courts in their efforts to draft a better instruction.

REFERENCES

- Bermant, G. (1986). Two conjectures about the issue of eyewitness testimony. *Law and Human Behavior*, 10, 97-100.
- Bibicoff, J. M. (1983). Seeing is believing? The need for cautionary jury instructions on the unreliability of eyewitness identification testimony. *San Fernando Valley Law Review*, 11, 95-122.
- Black, B. (1988). Evolving legal standards for the admissibility of scientific evidence. *Science*, 239, 1508-1512.
- 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.
- Brigham, J. C., & Bothwell, R. K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identification. *Law and Human Behavior*, 7, 19-30.
- Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own- and other-race persons. *Journal of Social Issues*, 41, 139-155.
- Buckhout, R. (1986). Personal values and expert testimony. *Law and Human Behavior*, 10, 127-144.
- Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1988, March). *The eyewitness, the expert psychologist, the judge and the jury*. Paper presented at the meeting of the American Psychology-Law Society, Miami, FL.
- 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.
- Deffenbacher, K. A., & Loftus, E. F. (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law and Human Behavior*, 6, 15-30.
- Ellsworth, P. C., & Getman, J. (1987). The use of social science in legal decision making. In L. Lipson and S. Wheeler (Eds.), *Law and the social sciences* (pp. 581-636). New York: Russell Sage.
- Federal Rules of Evidence* (1984). St. Paul, MN: West.
- Fox, S. G., & Walters, H. A. (1986). The impact of general versus specific expert testimony and eyewitness confidence upon mock juror judgment. *Law and Human Behavior*, 10, 215-228.
- Frazier, P., & Borgida, E. (1988). Juror common understanding and the admissibility of rape trauma syndrome evidence in court. *Law and Human Behavior*, 12, 101-122.
- Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).
- Giannelli, P. (1980). The admissibility of novel scientific evidence: *Frye vs. United States*, a half-century later. *Columbia Law Review*, 80, 1197-1250.
- Giannelli, P. (1983). *Frye vs. United States*. *Federal Rules Decisions*, 99, 189-202.
- Greene, E. (1988). Judge's instruction on eyewitness testimony: Evaluation and revision. *Journal of Applied Social Psychology*, 18, 252-276.
- Hastie, R. (1980). *From eyewitness testimony to beyond reasonable doubt*. Paper presented at the annual meeting of the Law and Society Association and the Research Committee on Sociology of Law of the International Sociological Association, Madison, WI.
- Hosch, H. M., Beck, E. L., & McIntyre, P. (1980). Influence of expert testimony regarding eyewitness accuracy on jury decisions. *Law and Human Behavior*, 4, 287-296.
- Konecni, V. J., & Ebbesen, E. B. (1986). Courtroom testimony by psychologists on eyewitness identification issues: Critical notes and reflections. *Law and Human Behavior*, 10, 117-126.
- Lempert, R. (1986). Social sciences in court: On "eyewitness experts" and other issues. *Law and Human Behavior*, 10, 167-181.
- 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.
- Lloyd-Bostock, S., & Clifford, B. R. (Eds.). (1983). *Evaluating witness evidence*. Chichester, England: Wiley.
- Loftus, E. F. (1983). Silence is not golden. *American Psychologist*, 65, 9-15.
- Loftus, E. F. (1986). Experimental psychologist as advocate or impartial educator. *Law and Human Behavior*, 10, 63-78.
- Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus." *Law and Human Behavior*, 11, 55-62.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 19-31.
- McCloskey, M., & Egeth, H. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist*, 38, 550-563.
- McCloskey, M., Egeth, H., & McKenna, J. (1986). The experimental psychologist in court: The ethics of expert testimony. *Law and Human Behavior*, 10, 1-13.
- Monahan, J., & Walker, L. (1988). Social science research in law: A new paradigm. *American Psychologist*, 43, 465-472.
- Pachella, R. G. (1986). Personal values and the value of expert testimony. *Law and Human Behavior*, 10, 145-150.
- Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139-156.
- State of Arizona v. Chapple*, 660 Pac.2d 1983.
- Tooley, V., Brigham, J. C., Maass, A., & Bothwell, R. K. (1987). Facial recognition: Weapon effect and attentional focus. *Journal of Applied Social Psychology*, 17, 845-859.
- United States v. Amaral*, 488 F.2d 1148 (9th Cir. 1973).
- United States v. Downing*, 753 F.2d (3rd Cir. 1985).
- United States v. Foshier*, 590 F.2d 381 (1st Cir. 1979).
- United States v. Smith*, 736 F.2d 1103 (6th Cir. 1984).
- United States v. Telfaire*, 469 F.2d 552 (D.C. Cir. 1972).
- Walker, L., & Monahan, J. (1987). Social frameworks: A new use of social science in law. *Virginia Law Review*, 73, 559-598.
- Wells, G. L. (1984). How adequate is human intuition for judging eyewitness testimony? In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 256-272). New York: Cambridge University Press.
- Wells, G. L. (1986). Expert psychological testimony: Empirical and conceptual analysis of effects. *Law and Human Behavior*, 10, 83-95.
- Wells, G. L., Lindsay, R. C. L., & Tournant, J. P. (1980). Effects of expert psychological advice on human performance in judging the validity of eyewitness testimony. *Law and Human Behavior*, 4, 275-285.
- Wells, G. L., & Loftus, E. F. (Eds.). (1984). *Eyewitness testimony: Psychological perspectives*. New York: Cambridge University Press.
- Wells, G. L., & Murray, D. M. (1984). Eyewitness confidence. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 155-170). New York: Cambridge University Press.
- Yarmey, A. D., & Jones, H. P. T. (1983). Is the psychology of eyewitness identification a matter of common sense? In S. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 13-40). Chichester, England: Wiley.