

Eyewitness Identification: Information Gain From Incriminating and Exonerating Behaviors

Gary L. Wells and Elizabeth A. Olson
Iowa State University

An information-gain approach to the analysis and interpretation of eyewitness identification data is described. The information-gain analysis is grounded in Bayesian statistics, permitting the important role of prior probabilities to be explored. This approach also forces a more complete treatment of the data and reveals important patterns that have escaped previous attention in the eyewitness identification literature. Particularly important is the ability of information-gain analyses to make salient the exonerating value of eyewitness behaviors rather than just their incriminating value. Analyses of sample data sets show how the exonerating value of filler identifications and "not there" responses can actually exceed the incriminating value of identifications of the suspect at certain points in the distribution of prior probabilities.

Following a crime for which there was one or more eyewitnesses, police investigators will often conduct a lineup. Each year in the United States, an estimated 77,000 people become criminal defendants after being identified from a lineup (Goldstein, Chance, & Schneller, 1989). In a lineup, a suspect is placed among fillers (sometimes called *distractors* or *foils*) who are known to be innocent of the offense. The eyewitness is then asked if one of the lineup members is the culprit. The task for the eyewitness seems simple at first glance, but mistaken identifications by eyewitnesses are responsible for more actual cases of wrongful conviction by juries than all other causes combined (see Scheck, Neufeld, & Dwyer, 2000; Wells et al., 1998). A large experimental literature shows that mistaken identification rates can be quite high under certain conditions and that eyewitness identification testimony can be very persuasive even when the identification is mistaken (Wells, 1993). Lab-based research on eyewitness identification has had an unprecedented impact on the legal system recently, especially with regard to how to make lineups yield more reliable identification evidence (Wells et al., 2000).

The purpose of the current article is to describe an information-gain approach to the analysis and understanding of eyewitness identification data. The information-gain approach uses Bayesian statistical logic, which we argue is very well suited to the problem at hand. We believe that this approach to eyewitness identification data will facilitate further progress in the literature in several ways. First, the information-gain approach helps maintain an important distinction between the direction of conditional probabilities ob-

tained in experimental research and the direction of conditional probabilities that evidence evaluators must consider. Experimental research estimates the probabilities of eyewitness behavior given the status of the suspect (suspect is the culprit vs. an innocent replacement). Evidence evaluators (e.g., detectives, prosecutors, judges, jurors), however, must estimate the probability of the suspect's status (culprit vs. innocent suspect) given the behavior of the eyewitness. Information-gain analyses help prevent confusion between these two very different probabilities. Second, the information-gain approach helps make clear the important real-world role played by prior probabilities (or base rates) despite their lack of representation in experiments. Third, the information-gain approach makes clear that some eyewitness responses to lineups are actually exonerating in terms of their directional meaning, a matter that has been nearly ignored in the eyewitness identification literature as well as in the real world. Finally, information-gain analyses permit direct comparisons of the informational value of incriminating versus exonerating eyewitness identification behaviors with some surprising conclusions about how these behaviors interact with the prior probabilities. Traditional inferential statistics are poorly suited for these purposes.

The distinction between the usual approach and our approach to eyewitness identification data parallels the distinction between the frequentist and the Bayesian approaches to data. Frequentist approaches using experimental designs focus on estimating the probability of an obtained set of data (behavior) given one condition versus some other condition. In an eyewitness experiment, for example, the traditional focus is on the probability that the eyewitness identifies the culprit given that the culprit is in the lineup versus identifies an innocent replacement for the culprit when the culprit is not in the lineup. The Bayesian approach focuses on the reverse of this conditional, namely, the probability the person is the culprit or not given that the eyewitness identified the person. The frequentist approach and the Bayesian approach yield the same probability only if the base rate for the lineup's containing the culprit is .50. Experiments typically set this base rate at .50 by making sure that equal numbers of participant witnesses view lineups that contain the culprit and do not contain the culprit. In

Gary L. Wells and Elizabeth A. Olson, Department of Psychology, Iowa State University.

This work was supported by National Science Foundation Grant SBR9807339 to Gary L. Wells. We thank Amy L. Bradfield, Steve Clark, Ron Fisher, Meg Gerrard, Richard Gonzalez, Rod Lindsay, Otto MacLin, Roy Malpass, Barry Scheck, and John Turtle for comments on drafts and oral presentations of this work.

Correspondence concerning this article should be addressed to Gary L. Wells, Department of Psychology, Iowa State University, W112 Lagomarcino Hall, Ames, Iowa 50011. E-mail: glwells@iastate.edu

generalizing to actual lineups conducted by police detectives, this is tantamount to assuming the suspect is the actual culprit in 50% of lineups and the detectives have the "wrong guy" the other 50% of the time. No one actually knows the real-world base rate for culprit-present lineups in any given jurisdiction. Although hypothesis testing often can be conducted independently of base rates, ecological validity sometimes requires considerations of base rates (see Dawes, 1996). Because of the profound impact of the base rate variable on eyewitness identification, it is important to use statistical tools that simulate variations in the base rate across all possible values.

The terms *prior probability* and *base rate* are used interchangeably in this article. Both refer to the prelineup chances that the lineup's suspect is the actual culprit. The term *prior probability* might make the most sense when thinking about a given, single suspect, in which case one might think about prior probability in terms of the amount of evidence that police have against the suspect prior to conducting the lineup. The term *base rate* might make the most sense when thinking of the problem across a large number of lineups and considering the percentage of times that the lineup's suspect is the culprit in question versus an innocent person.

In its simplest form, information gain is the difference between the probability that the suspect is the culprit before conducting the lineup (a prior probability) and the probability that the suspect is the culprit after conducting the lineup (a posterior probability). Because base rates or prior probabilities are necessary for calculating information gain, this problem is well suited to a Bayesian statistical analysis. This is not the first article to use Bayesian statistics for the analysis of eyewitness identification data. Bayesian analyses of eyewitness identification data were first used by Wells and Lindsay (1980) and were later used by Wells and Turtle (1986) to analyze specific hypotheses about how lineups reduce uncertainty regarding the guilt or innocence of a criminal suspect. Levi (1998) used Bayesian statistics to examine posterior probabilities for given likelihood ratios across a set of possible prior probabilities for a given eyewitness behavior, namely, the identification of a suspect. The current article goes much further. We propose using information gain as a way to characterize eyewitness identification data; we describe the importance of examining all four possible eyewitness behaviors; we describe tests of statistical significance; we show how incriminating responses and exonerating responses have information-gain curves with different skews, which guarantees that their curves will intersect; and we describe the meaning of these intersections.

Critical Concepts and Definitions of Terms

It is important to maintain a clear understanding of the meaning of several terms. The term *culprit* refers to the guilty party, the person who the witness saw commit the offense. The term *suspect* (S) refers to someone who might or might not be the culprit; the suspect could be an innocent suspect (S not culprit) or the suspect could be the culprit (S is culprit). The term *filler* refers to someone who is used as a stand-in or foil in the lineup. Fillers are not suspects and instead are "known innocents." The identification of a filler by an eyewitness does not result in the identified person's being charged with a crime. Although the identification of a filler is an error, it is a "known error." The identification of a suspect, however, might or might not be a mistaken identification, depending on whether the suspect is the culprit or not. Throughout this article, we assume that a lineup contains only one suspect, who might or might not be the culprit, and the remaining members are known-innocent fillers. The issues, both mathematically and logically, change dramatically when a lineup has multiple suspects (Wells & Turtle, 1986). Under the assumption that a lineup has only one suspect, a mistaken identification of a suspect cannot occur when the culprit is in the lineup (because the suspect is the culprit); any misidentification would have to be a filler identification. The single-suspect assumption also means that the terms *culprit-present lineup* and *culprit-absent lineup* are synonymous with the phrases "suspect is culprit" and "suspect not culprit", respectively. It is also important to recognize that there are four possible behaviors or responses to a lineup that an eyewitness can make: The eyewitness can identify the suspect (IDS), the eyewitness can identify a filler (IDfiller), the eyewitness can indicate that the culprit is not in the lineup ("not there"), or the eyewitness can say "don't know."

As shown in Table 1, the four possible behaviors can be combined with the two states of truth (S is the culprit or not) to yield eight possible outcomes. The labels *hit*, *mistaken identification*, *miss*, and *correct rejection* are well-known, traditional labels given to these outcomes in signal detection theory (Swets, Tanner, & Birdsall, 1961). Notice that we do not call the identification of a filler a mistaken identification. The identification of a filler is a mistake, but it is not the type of mistake that results in charges being brought against the filler. By definition, fillers are known a priori to be innocent of the offense in question. Hence, we reserve the term *mistaken identification* for instances in which the eyewitness identifies the innocent suspect. "Don't know" responses have different implications when the suspect is the culprit than when the

Table 1
Eight Possible Lineup Outcomes

State of truth	Witness behavior			
	Identification of suspect	Identification of filler	"Not there" response	"Don't know" response
Suspect is culprit	Hit: $p(\text{IDS} \text{S is culprit})$	Filler ID: $p(\text{IDfiller} \text{S is culprit})$	Miss: $p(\text{"not there"} \text{S is culprit})$	Miss, Type 2: $p(\text{"don't know"} \text{S is culprit})$
Suspect is not culprit	Mistaken ID: $p(\text{IDS} \text{S not culprit})$	Filler ID: $p(\text{IDfiller} \text{S not culprit})$	Correct rejection: $p(\text{"not there"} \text{S not culprit})$	Correct rejection, Type 2: $p(\text{"don't know"} \text{S not culprit})$

Note. ID = identification; IDS = identification of suspect; S = suspect; IDfiller = identification of filler.

suspect is not the culprit. We consider “don’t know” responses to be a type of *miss* when the suspect is the culprit and a type of *correct rejection* when the suspect is not the culprit.

In actual cases (unlike experiments), the behavior of the eyewitness is known but the state of truth (whether S is the culprit or not) is unknown. All that is known in actual cases is that the eyewitness has engaged in one of the four possible behaviors. The state of truth (S is or is not the culprit) is the variable we are trying to predict from the behavior of the eyewitness. For each possible behavior (e.g., identification of a filler), the critical data for our purposes concern how often that behavior occurs when the suspect is the culprit versus when the suspect is not the culprit. These values are estimated by the conditional probability expressions in Table 1. For example, the conditional probability value $p(\text{IDS}|S \text{ is culprit})$ is the probability of an identification of the suspect given that the suspect is the culprit, $p(\text{IDS}|S \text{ not culprit})$ is the probability of an identification of the suspect given that the suspect is not the culprit, $p(\text{IDfiller}|S \text{ is culprit})$ is the probability of a filler identification given that the suspect is the culprit, and so on. Notice the conditional probabilities of the behaviors can be calculated directly from an experiment. However, the posterior probabilities that the suspect is the culprit, for example, $p(S \text{ is culprit}|\text{IDS})$, cannot be calculated directly from an experiment. The latter values, which are necessary for estimating information gain, require consideration of the prior probability value $p(S \text{ is culprit})$.

Posterior Probabilities and the Role of Priors

Using the above terminology, let $p(S \text{ is culprit})$ be the *prior* (prelineup) probability that the suspect is the culprit, $p(S \text{ not culprit})$ be the prior probability that the suspect is not the culprit, $p(S \text{ is culprit}|\text{IDS})$ be the *posterior* (postlineup) probability that the suspect is the culprit given that the eyewitness identified the suspect, $p(S \text{ is culprit}|\text{IDfiller})$ be the posterior probability that the suspect is the culprit given that the eyewitness identified a filler, $p(S \text{ is culprit}|\text{“not there”})$ be the posterior probability that the suspect is the culprit given that the eyewitness said “not there,” and $p(S \text{ is culprit}|\text{“don’t know”})$ be the posterior probability that the suspect is the culprit given that the eyewitness said “don’t know.”

Bayes’s theorem, and its simple algebraic variations, permits us to specify a number of useful equations relating these parameters.

In particular, we are interested in estimating the posterior probabilities that the suspect is the culprit given one of the behaviors of the eyewitness. Equations 1–4 (presented in Table 2) are posterior probabilities, each reflecting a different value depending on the behavior of the eyewitness. To estimate each of these posterior probabilities, three parameters have to be estimated. In the case of $p(S \text{ is culprit}|\text{IDS})$, one must estimate $p(\text{IDS}|S \text{ is culprit})$, $p(\text{IDS}|S \text{ not culprit})$, and $p(S \text{ is culprit})$. The parameter $p(S \text{ is culprit})$ is simply $1 - p(S \text{ is culprit})$ and therefore is not a separate parameter. Two of these three parameters, namely, $p(\text{IDS}|S \text{ is culprit})$ and $p(\text{IDS}|S \text{ not culprit})$, are estimable for any condition of an eyewitness experiment (see Table 1). The ratio of the these two parameters [$p(\text{IDS}|S \text{ is culprit})/p(\text{IDS}|S \text{ not culprit})$] in Equation 1 (see Table 2) is known as the *likelihood ratio* in Bayes’s theorem. In effect, the likelihood ratio reflects how much more often the behavior (e.g., identification of suspect) occurs when the suspect is guilty than when the suspect is innocent.

The prior probability parameter $p(S \text{ is culprit})$, however, cannot be estimated from an experiment. In a typical experiment, $p(S \text{ is culprit})$ is arbitrarily set at .50: Half the eyewitnesses are shown a lineup in which the suspect is the culprit, and half are shown a lineup in which the suspect is replaced with an innocent person. In actual cases, the base-rate frequency with which eyewitnesses confront lineups with a guilty suspect, that is, $p(S \text{ is culprit})$, relative to the base-rate frequency with which eyewitnesses confront lineups that do not include the guilty suspect, that is, $p(S \text{ not culprit})$, is unknown. Furthermore, this base rate is likely to vary from one jurisdiction to another depending on the amount of evidence that police gather on that suspect before conducting the lineup (Wells, 1993).

Whether the lineup includes the actual culprit or not is a powerful variable affecting the behavior of the eyewitness (Wells & Lindsay, 1980). Errors of commission (identifying an innocent suspect as well as identifying fillers) are more common when the actual culprit is not in the lineup than when the culprit is in the lineup. “Not there” responses are also more likely when the actual culprit is not in the lineup than when the culprit is in the lineup. Correct identifications, of course, can occur only when the culprit is in the lineup, and identifications of an innocent suspect can occur only when the culprit is not in the lineup. The status of

Table 2
Equations for Posterior Probability That the Suspect Is the Culprit Given Eyewitness Behavior

Posterior probability	Equation
1. $p(S \text{ is culprit} \text{IDS})$	$\frac{p(\text{IDS} S \text{ is culprit}) p(S \text{ is culprit})}{p(\text{IDS} S \text{ is culprit}) p(S \text{ is culprit}) + p(\text{IDS} S \text{ not culprit}) p(S \text{ not culprit})}$
2. $p(S \text{ is culprit} \text{IDfiller})$	$\frac{p(\text{IDfiller} S \text{ is culprit}) p(S \text{ is culprit})}{p(\text{IDfiller} S \text{ is culprit}) p(S \text{ is culprit}) + p(\text{IDfiller} S \text{ not culprit}) p(S \text{ not culprit})}$
3. $p(S \text{ is culprit} \text{“not there”})$	$\frac{p(\text{“not there”} S \text{ is culprit}) p(S \text{ is culprit})}{p(\text{“not there”} S \text{ is culprit}) p(S \text{ is culprit}) + p(\text{“not there”} S \text{ not culprit}) p(S \text{ not culprit})}$
4. $p(S \text{ is culprit} \text{“don’t know”})$	$\frac{p(\text{“don’t know”} S \text{ is culprit}) p(S \text{ is culprit})}{p(\text{“don’t know”} S \text{ is culprit}) p(S \text{ is culprit}) + p(\text{“don’t know”} S \text{ not culprit}) p(S \text{ not culprit})}$

Note. S = suspect; IDS = identification of suspect; IDfiller = identification of filler.

“don’t know” responses is less clear; “don’t know” responses might occur equally often when the culprit is in the lineup as when the culprit is not in the lineup. We report the first analysis in the literature on the issue of whether “don’t know” responses have informational value. Because of the powerful impact of the base-rate parameter $p(S \text{ is culprit})$ on eyewitness behavior, the important mathematical effect of $p(S \text{ is culprit})$ on the posterior probabilities, and the fact that $p(S \text{ is culprit})$ is unknown in actual cases, posterior probabilities in our analyses are calculated across the entire range of possible values of the parameter $p(S \text{ is culprit})$ from 0.00 to 1.00.

As an example, consider how we could use the values of $p(\text{IDS}|S \text{ is culprit})$ and $p(\text{IDS}|S \text{ not culprit})$ to calculate the posterior probability $p(S \text{ is culprit}|\text{IDS})$ for different values of the prior probability $p(S \text{ is culprit})$ variable. Consider an example in which $p(\text{IDS}|S \text{ is culprit})$ is .60 and $p(\text{IDS}|S \text{ not culprit})$ is .15. If the prior probability $p(S \text{ is culprit})$ is .80, then it follows from Equation 1 that $p(S \text{ is culprit}|\text{IDS}) = .94$. Suppose, however, the prior probability $p(S \text{ is culprit})$ were changed to .20, while holding constant the values of $p(\text{IDS}|S \text{ is culprit})$ and $p(\text{IDS}|S \text{ not culprit})$ at .60 and .15, respectively. When $p(S \text{ is culprit})$ is .20, then $p(S \text{ is culprit}|\text{IDS}) = .50$. In these two examples, we have held constant the recognition memory capabilities of the eyewitnesses and varied only the prior probability $p(S \text{ is culprit})$. Notice, however, that the chances that the identification was mistaken changed from a mere 6% ($1.00 - .94$) to 50% ($1.00 - .50$). This simple example makes it clear that the probability that an identified suspect is the actual culprit depends not only on the capabilities of eyewitnesses (as calculated in experiments and captured by the likelihood ratio) but also on the base rate of the actual culprit’s being in the lineup, which is $p(S \text{ is culprit})$.

Researchers have argued that people neglect base-rate information in making a variety of judgments (Kahneman & Tversky, 1973), but actual usage depends on task structure and representation (Koehler, 1996). The eyewitness identification literature has tended to be base-rate neglectful in the sense that the base-rate parameter $p(S \text{ is culprit})$ is set at a fixed value of .50 in experiments. Although the base-rate value of $p(S \text{ is culprit})$ is of no particular relevance for most hypothesis tests in eyewitness identification experiments, it is critical when attempting to estimate the probability that a suspect is the culprit based on whether or not the eyewitness identifies the suspect. For purposes of our analysis, $p(S \text{ is culprit})$ is a variable, not a constant.

We can draw an analogy between the role of prior probability in police lineups and the role of prior probability in other diagnostic settings, such as medical diagnoses (Eddy, 1982). For example, suppose we were interested in the probability that someone has disease D given that he or she displays symptom A . Suppose medical data indicate the probability of symptom A given disease X is .90, and the probability of symptom A given that the person does not have disease D is .15. What is the probability that the person has disease D ? This cannot be calculated without knowing the base rate for disease D . If the base rate for disease D is low (e.g., 1% of the population), then the posterior probability that the person has disease D is still rather small (less than 6%). If the base rate for disease D is higher (e.g., 15% of the population), then the posterior probability that the person has disease D is considerably larger (about 35%). The problem of prior probabilities is not unique to lineups and, in spite of the difficulty that some people

might have in forming a clear mental representation of prior probabilities, they have real consequences for the correctness of our conclusions about the probabilities of events.

Information Gain

The amount of information gained from a given eyewitness who makes a given response to a lineup is indexed by the absolute difference between the prior probability $p(S \text{ is culprit})$ and the posterior probability $p(S \text{ is culprit}|\text{IDS})$; Wells & Lindsay, 1980). For example, if $p(S \text{ is culprit}) = .80$ and $p(S \text{ is culprit}|\text{IDS}) = .94$ (as in the first numerical example), then information gain from an identification of the suspect is .14. If $p(S \text{ is culprit}) = .20$ and $p(S \text{ is culprit}|\text{IDS})$ is .50 (as in the second numerical example), then information gain from an identification of the suspect is .30. The information-gain formula for an identification of the suspect is

$$|p(S \text{ is culprit}) - p(S \text{ is culprit}|\text{IDS})|.$$

Information gain for a filler identification is

$$|p(S \text{ is culprit}) - p(S \text{ is culprit}|\text{IDfiller})|.$$

Information gain for a “not there” response is

$$|p(S \text{ is culprit}) - p(S \text{ is culprit}|\text{“not there”})|.$$

Information gain for a “don’t know” response is

$$|p(S \text{ is culprit}) - p(S \text{ is culprit}|\text{“don’t know”})|.$$

Note that information gain is expressed as an absolute value rather than a directional value. For incriminating witness behaviors (identifications of the suspect), the prior will be lower than the posterior; for exonerating behaviors (e.g., “not there” responses), the prior will be higher than the posterior. Using absolute values for the prior–posterior difference permits direct comparisons of the informational value of different eyewitness behaviors regardless of the informational direction.

Information-Gain Curves

Any given eyewitness behavior (e.g., identifying the suspect) has a varying amount of information gain associated with it depending on the prior probability. Consider the prior probability $p(S \text{ is culprit})$ to be a continuous variable ranging from 0.00 to 1.00. Consider the likelihood ratio 4.0 from the above numerical example where $p(\text{IDS}|S \text{ is culprit}) = .60$ and $p(\text{IDS}|S \text{ not culprit}) = .15$. Figure 1 shows an information-gain curve, which plots the prior probability $p(S \text{ is culprit})$ on the x -axis and information gain on the y -axis. Of course, there is no information gain (prior probability – posterior probability = 0) when the prior probability $p(S \text{ is culprit})$ is either 0.00 or 1.00.

These information-gain curves, first used by Wells and Lindsay (1980), are readily translated into a posterior probability at any given prior probability value. For identifications of the suspect, the posterior probability $p(S \text{ is culprit}|\text{IDS})$ is simply the sum of the prior probability $p(S \text{ is culprit})$ and the information gain associated with that prior. For exonerating behaviors, such as “not there” responses, the posterior probability is the difference between the prior probability and information gain.

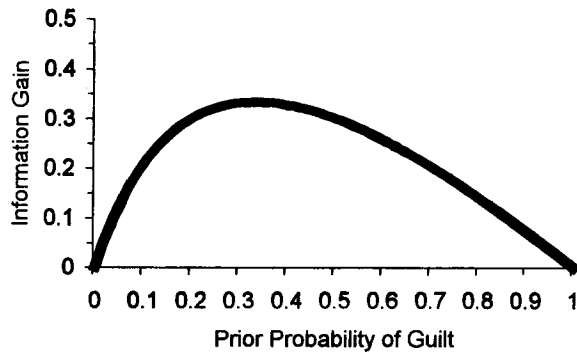


Figure 1. Hypothetical information-gain curve.

Sample Analyses of Past Experiments

We illustrate the information-gain approach to eyewitness identification by using data from some previously published experiments. It is not our intent to present these data as representative of the eyewitness identification literature. Eyewitness identification experiments vary considerably along a number of dimensions such as (a) how good a view the participant witnesses were given of the culprit, (b) the length of time between witnessing and the presentation of the lineup, (c) the size of the lineup, (d) instructions to witnesses prior to viewing the lineup, and so on. Nor is it our intent to conduct a meta-analysis. Instead, we use data from some past experiments to show how information gain from one type of eyewitness behavior (e.g., identification of suspects) can be compared with information gain from another type of behavior (e.g., filler identifications) and the important role played by base rates. In addition, these analyses show how the heights and skews of information-gain curves vary across conditions and how these curves can be interpreted. In the process of describing the analyses, we show how some regularities in the data, which can go unnoticed using traditional approaches, surface vividly and naturally using the information-gain approach.

The eyewitness identification literature has placed a premium on a particular class of variables called *system variables* because system variables are under control of the justice system and affect the accuracy of eyewitness identifications (Wells, 1978). Two of the most prominent system variables are sequential versus simultaneous lineup presentation procedures and methods for selecting fillers in lineups. We illustrate our analyses using some previously published experiments that entailed manipulation of either the sequential versus simultaneous variable or methods for selecting lineup fillers. This permitted us to not only compare the information gain for one type of eyewitness behavior with another but also to compare information gain for each behavior under one set of conditions with that under another set of conditions. In all the experiments, the recommended method of instructing the eyewitnesses that the actual culprit might or might not be in the lineup was used (Malpass & Devine, 1981).

Simultaneous Versus Sequential Data

Information-gain analyses require large enough sample sizes to ensure sufficient numbers of any relatively rare witness behaviors to calculate conditional probabilities for those behaviors. The

sequential lineup presents a special challenge because the frequency of filler identifications is relatively low. Accordingly, for the simultaneous versus sequential analyses, we collapsed the data over three separate experiments to secure a total sample size of 452. These data were obtained from articles by Lindsay, Lea, and Fulford (1991); Lindsay and Wells (1985); and Melara, DeWitt-Rickards, and O'Brien (1989). All three studies used very similar methodologies involving culprit-present and culprit-absent lineups, lineup members were presented either simultaneously or one at a time, witnesses were instructed that the culprit might or might not be in the lineup, and the lineup fillers all fit a general verbal description of the culprit. Before collapsing the data across the three experiments, an initial test was conducted to see if relative choosing rates in the culprit-present and culprit-absent conditions were similar across the three experiments. The ratio of choosing in culprit-present versus culprit-absent lineups is a general index of performance (where a ratio of 1.00 indicates chance performance). Overall ratios for the Lindsay et al., Lindsay and Wells, and Melara et al. studies were quite similar (1.89, 1.37, and 1.40, respectively) and were not significantly different using chi-square tests (maximum chi square for the three single degree of freedom contrasts = 1.63, $p \cong .20$).

Filler-Selection Method Data

Data for examining the effects of filler-selection methods on information gain were obtained from Wells, Rydell, and Seelau (1993). In the Wells et al. study, all lineups were conducted using the simultaneous method. Thus, unlike the sequential studies, filler identifications were not rare behaviors, and a smaller total sample size (252) was sufficient for conducting the information-gain analyses. In the Wells et al. study, witnesses were exposed to live staged crimes committed by various crime actors. Lineups that included the culprit or an innocent replacement for the culprit were then constructed individually for each of 252 eyewitnesses using various methods for selecting lineup fillers. For purposes of illustrating information gain, we analyzed conditions in which fillers were selected using the mismatch-description method and the match-description method. The *mismatch-description method* represents the classic biased lineup: The suspect fits the verbal description of the culprit (which the eyewitness had given after the crime), but the fillers clearly do not fit that description. The *match-description method*, in contrast, uses both a suspect and fillers who fit the description of the culprit. As expected, the mismatch-description method yielded a high rate of identifications of the culprit (in culprit-present lineups) but yielded a high rate of identifications of the innocent suspect (in culprit-absent lineups). The match-description method yielded a similarly high rate of identifications of the culprit (in culprit-present lineups) but yielded a low rate of identifications of the innocent suspect (in culprit-absent lineups). Other research shows similar patterns for the match-description and mismatch-description methods of selecting fillers (e.g., Juslin, Olson, & Winman, 1996; Lindsay, Martin, & Webber, 1994; Tunnicliff & Clark, 2000).

Overview of Analyses

Our analyses of information gain are divided into four main sections. First, we examine information gain for simultaneous

versus sequential presentations. These information-gain curves are organized by the behavior of the eyewitness. Specifically, we first compare the simultaneous-sequential information-gain curves for identifications of the suspect, then for "not there" responses, and then for filler identifications. There are no information-gain analyses for "don't know" responses for the simultaneous versus sequential studies because these studies have not separated "don't know" responses from "not there" reactions. In the second section, we examine information gain for the two filler-selection methods, namely, the match-description method and the mismatch-description method. Again, the information-gain curves are organized by the behavior of the eyewitness. Specifically, we first examine the effect of filler-selection methods on information gain for identifications of the suspect, then for "not there" responses, then for filler identifications, and then for "don't know" responses. In the third section, we compare information gain for each eyewitness behavior with each other behavior for both the simultaneous versus sequential data and for the filler-selection method data. In the fourth section, we describe how to conduct significance tests for determining whether an information-gain curve is greater than zero or whether one information-gain curve is significantly greater than some other information-gain curve.

Simultaneous Versus Sequential Presentations

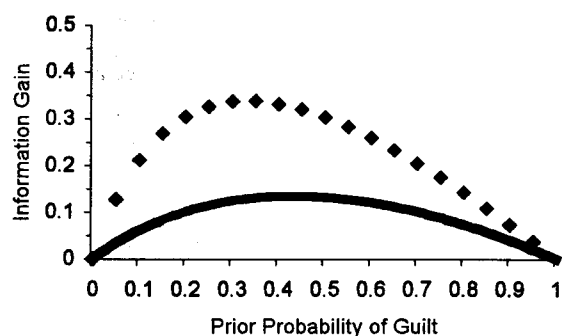
We first analyzed information gain as a function of whether the lineup was presented simultaneously or sequentially. The typical police lineup is a simultaneous lineup in which the suspect is embedded among fillers and all lineup members are shown to the eyewitness at once. The eyewitness is then asked to decide if the actual culprit is in the lineup and, if so, which member of the lineup is the culprit. The sequential lineup, in contrast, shows the eyewitness each lineup member one at a time and requires the eyewitness to make a decision on each one (that is the culprit or that is not the culprit) before viewing the remainder. First tested in 1985 (Lindsay & Wells, 1985), the sequential presentation method has shown a clear and consistent advantage over the simultaneous method (e.g., Cutler & Penrod, 1998; Lindsay et al., 1991; Sporer, 1993). When the lineup contains the actual culprit (i.e., when the suspect is the culprit), both the simultaneous and the sequential methods yield relatively similar rates of identifying the culprit. When the lineup does not contain the actual culprit (i.e., the suspect is not the culprit), however, the simultaneous method yields a higher rate of mistaken identifications than does the sequential method. The ability of the sequential lineup procedure to outperform the simultaneous lineup procedure is known as the *sequential-superiority effect*.

It may come as no surprise that information gain for identifications of the suspect may be higher for the sequential than for the simultaneous procedure because previous tests of the sequential-superiority effect have relied on identifications of the suspect to reach the conclusion that the sequential procedure is superior to the simultaneous procedure. However, previous analyses have not examined how the sequential versus simultaneous procedures affect information gain for identifications of fillers and for nonidentifications. Furthermore, previous analyses have not examined how the prior probability variable $p(S \text{ is culprit})$ moderates the relative advantage of the sequential over the simultaneous procedure.

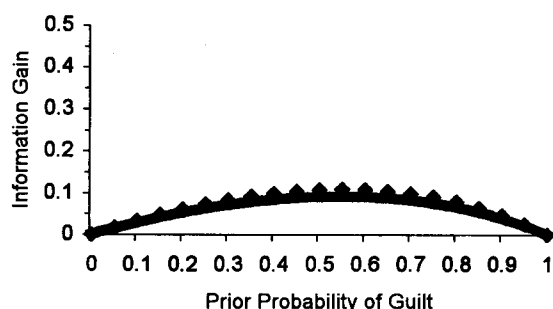
Information gain for identifications of the suspect. First, we calculated information gain from identifications of the suspect separately for simultaneous and sequential lineups. We used averages weighted by sample sizes to collapse the data across the three experiments. For the simultaneous lineup procedure, the average rate of selecting the suspect when the suspect was the culprit was .55, and the average rate of selecting the suspect when the suspect was not the culprit was .32. For the sequential lineup procedure, the average rate of selecting the suspect when the suspect was the culprit was .47, and the average rate of selecting the suspect when the suspect was not the culprit was .12. Calculations of posterior probabilities $p(S \text{ is culprit} | \text{IDS})$ were made using these values for all possible values of the prior probability $p(S \text{ is culprit})$, and the difference score (information gain) was graphed across all priors. The result is shown in Figure 2A. Notice the area under the curve is greater for the sequential than for the simultaneous lineup procedure, as a result expected from what was already known about the sequential-superiority effect. Note as well the difference in the height of the curves diminishes at the extremes of the priors. This follows from the simple principle that new information (the behavior of the eyewitness) has less impact when prior uncertainty is low (i.e., when the prior probability is close to 1.00 or 0.00 rather than near maximum uncertainty of .50).

Both the information-gain curves in Figure 2A are skewed to the right, which results in peak information gain when the prior probability $p(S \text{ is culprit})$ was below .50. This direction of skew follows from the principle that new information has more impact if it goes against the prior probability than if it is in the same direction as the prior probability. In the case of identifications of the suspect, the maximum information gain for the sequential lineup occurred when the prior probability $p(S \text{ is culprit})$ was .33, at which point information gain was .34. The maximum information gain for the simultaneous lineup peaked when the prior probability $p(S \text{ is culprit})$ was .43, at which point information gain was .14. The maximum advantage of the sequential over the simultaneous lineup (where the two information-gain curves show the greatest difference) occurred at a prior probability $p(S \text{ is culprit})$ of .28 (at which point information gain was .20 higher for the sequential than for the simultaneous lineup). The interpretation of this result is rather straightforward: Greater information gain from identifications of the suspect for the sequential lineup means the chances that the suspect was innocent were lower when the sequential lineup procedure was used than when the simultaneous procedure was used.

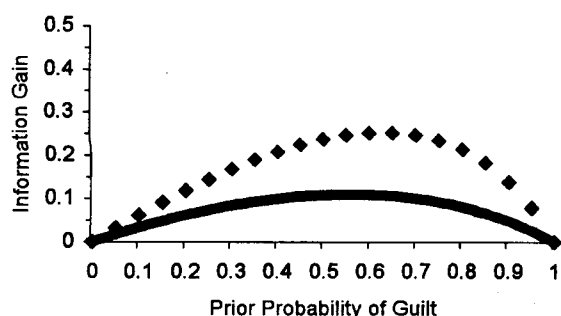
Information gain for "not there" responses. For the simultaneous lineup procedure, the average rate of "not there" responses was .40 when the suspect was not the culprit and .26 when the suspect was the culprit. For the sequential lineup procedure, the average rate of "not there" responses was .72 when the suspect was not the culprit and .47 when the suspect was the culprit. Notice these two likelihood ratios are nearly identical (i.e., 40:26 \cong 73:47 \cong 1.5:1). Hence, it should come as no surprise that information gain from a "not there" decision was approximately the same for the simultaneous as for the sequential lineup as shown in Figure 2B. Both information-gain curves peaked when $p(S \text{ is culprit})$ was .55. The nearly equivalent information gain from "not there" responses for the sequential and simultaneous lineups means that the chances that the suspect is innocent or guilty when the eye-



A



B



C

Figure 2. A: Information gain from identifications of suspect for the simultaneous versus sequential lineup. B: Information gain from "not there" decisions for the simultaneous versus sequential lineup. C: Information gain from filler identifications for the simultaneous versus sequential lineup. Solid lines represent simultaneous lineups; solid squares represent sequential lineups.

witness says "not there" are approximately the same for sequential and simultaneous lineups.

Notice that, unlike those for identifications of the suspect, information-gain curves for "not there" decisions are skewed to the left. Accordingly, the maximum information gain for "not there" responses occurred when the value of $p(S \text{ is culprit})$ was higher

than .50. This is precisely what would be expected from the fact that "not there" decisions are exculpatory (more likely to be associated with lineups in which the suspect is not the culprit than with lineups in which the suspect is the culprit) rather than incriminatory.

Information gain for filler identifications. For the simultaneous lineup procedure, the average rate of filler identifications was .29 when the suspect was not the culprit and .19 when the suspect was the culprit. For the sequential lineup procedure, the average rate of filler identifications was .16 when the suspect was not the culprit and .06 when the suspect was the culprit. Notice that the likelihood ratio was higher for the sequential ($16:6 \cong 2.7$) than for the simultaneous lineup procedure ($29:19 \cong 1.5$). This difference is reflected in the information-gain curves shown in Figure 2C. Information gain for the simultaneous lineup procedure peaked at .11 when $p(\text{suspect is culprit})$ was .55. Information gain for the sequential lineup procedure peaked at .25 when $p(S \text{ is culprit})$ was .63. Greater information gain from filler identifications for the sequential lineup means that the chances that the suspect is innocent are higher when a filler identification occurs with a sequential lineup procedure than when a filler identification occurs with a simultaneous lineup.

As with those for "not there" decisions, information-gain curves for filler identifications are skewed to the left and maximum information gain occurs when $p(S \text{ is culprit})$ was greater than .50. Again, this is what should happen when the behavior (in this case a filler identification) is exculpatory rather than incriminatory.

Summary. The greater information gain associated with the sequential over the simultaneous lineup occurred primarily for identifications of the suspect and, to a lesser extent, filler identifications. "Not there" responses had relatively low information gain overall and showed no differences in information gain between sequential and simultaneous lineups.

Filler-Selection Methods

The filler-selection problem is one of the oldest in the modern eyewitness identification literature: How should criminal investigators select fillers to use in the lineup so that the fillers adequately protect the suspect? Early treatments recommended selecting fillers who fit the eyewitness's initial verbal description of the culprit (Doob & Kirshenbaum, 1973; Wells, Leippe, & Ostrom, 1979). As mentioned above, we call this the *match-description filler-selection method*. As long as the innocent suspect is simply someone who fits the description of the culprit, the match-description method for fillers assures there is only a $1/K$ (where K is the number of lineup members) probability that an eyewitness who identifies someone will identify the innocent suspect (the remaining probability is distributed across the fillers; see Luus & Wells, 1991). In contrast, consider a method for selecting fillers in which the suspect matches the description of the culprit but the fillers do not. Experimental data show that using fillers who do not fit the description of the culprit leads eyewitnesses to select the suspect who fits the description even when the suspect is innocent (Lindsay & Wells, 1980).

The information-gain comparison of the two filler-selection methods is interesting for two primary reasons. First, the effect of filler-selection methods has never been examined on information gain for filler identifications and "not there" responses. Will the

superiority of the match-description method be reflected in information gain for those behaviors as well? Second, filler-selection method studies have partitioned nonidentification responses separately into "not there" responses and "don't know" responses. The simultaneous versus sequential experiments, however, do not include this breakdown of nonidentifications. Logically, we might expect "don't know" responses to have little or no information gain, but this has not been tested. Possibly, "don't know" responses are more likely when the suspect is not the culprit than when the suspect is the culprit, which translates into information gain. If "don't know" responses have low informational value, however, partitioning out those who say "don't know" from those who explicitly say "not there" should increase the informational value of "not there" responses.

Information gain for identifications of the suspect. First, we calculated information gain from identifications of the suspect separately for each of the two filler-selection methods. For the match-description method, the average rate of selecting the suspect when the suspect was the culprit was .67, and the average rate of selecting the suspect when the suspect was not the culprit was .12. For the mismatch-suspect method, the average rate of selecting the suspect when the suspect was the culprit was .72, and the average rate of selecting the suspect when the suspect was not the culprit was .46. Calculations of the posterior probabilities $p(S \text{ is culprit} | IDS)$ were made using these values for all possible values of the prior probability $p(S \text{ is culprit})$, and the resulting information-gain scores were graphed across all priors. The result is shown in Figure 3A. As expected, information gain for identifications of the suspect was greater for the match-description method than for the mismatch-description method. Greater information gain for the match-description filler-selection method means the chance that an identified suspect is guilty is greater with

the match-description method than with the mismatch-description method.

Note as well that peak information gain occurred at a relatively low level of the prior probability variable $p(S \text{ is culprit})$. Again, this was expected because new information has more impact if it goes against the prior probability than if it is in the same direction as the prior probability. Maximum information gain for identifications of the suspect with the match-description method occurred when the prior probability $p(S \text{ is culprit}) = .30$ (at which point information gain was .41). Maximum information gain for identifications of the suspect with the mismatch-description method occurred when the prior probability $p(S \text{ is culprit}) = .44$ (at which point information gain was .11).

Information gain for "not there" responses. Information gain for "not there" responses was calculated for both filler-selection methods. For the match-description filler-selection method, the average rate of "not there" responses was .32 when the suspect was not the culprit and .12 when the suspect was the culprit. For the mismatch-description filler-selection method, the average rate of "not there" responses was .32 when the suspect was not the culprit and .15 when the suspect was the culprit. Information-gain curves for each of the filler-selection methods are shown in Figure 3B.

Notice that the match-description and the mismatch-description conditions yielded similar levels of information gain for "not there" responses. Notice as well that, unlike those for identifications of the suspect, the information-gain curves for "not there" responses were skewed to the left for both the match-description and the mismatch-description curves, resulting in peak information gain when the prior probability was greater than .50. This is expected because new information (a "not there" response in this case) has more impact when it goes against the prior probability. Maximum information gain for the match-description method oc-

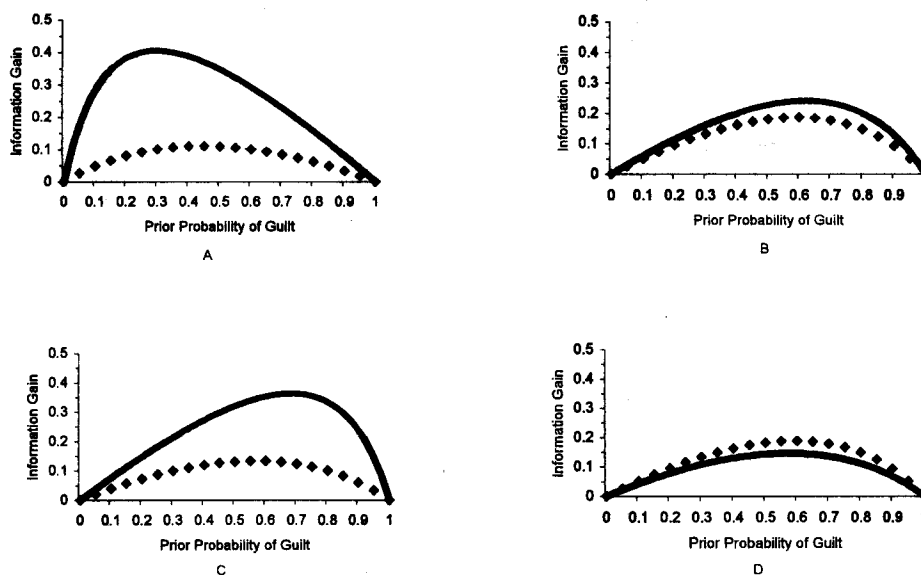


Figure 3. A: Information gain from identification of suspect as a function of filler-selection strategy. B: Information gain from "not there" responses as a function of filler-selection strategy. C: Information gain from filler identifications as a function of filler-selection strategy. D: Information gain from "don't know" responses as a function of filler-selection strategy. Solid lines represent match description; solid squares represent mismatch description.

curred when the prior probability $p(S \text{ is culprit})$ was .62 (at which point information gain was .24). Maximum information gain for the mismatch-description method occurred when the prior probability $p(S \text{ is culprit})$ was .59 (at which point information gain was .19).

Information gain for filler identifications. Information gain for filler identifications was calculated for both of the filler-selection methods. For the match-description method, the average rate of filler identifications was .32 when the suspect was not the culprit and .07 when the suspect was the culprit. For the mismatch-description method, the average rate of filler identifications was .12 when the suspect was not the culprit and .07 when the suspect was the culprit. Information-gain curves for filler identifications for both filler-selection methods are shown in Figure 3C.

Clearly, the match-description method produced more information gain for filler identifications than did the mismatch-description method. The low informational value of filler identifications with the mismatch-description method was the result of a low rate of filler identifications when the actual culprit was not in the lineup. This makes sense because none of the fillers in the mismatch-description method fit the description of the culprit and, hence, almost all identifications were instead flowing to the innocent suspect (who does fit the description).

Maximum information gain for filler identifications with the match-description method occurred when the prior probability $p(S \text{ is culprit}) = .68$ (at which point information gain was .36). Recall that in this case, we should shift our posterior probability that the suspect is the culprit downward: The prior probability .68 minus information gain .36 leaves us with a posterior probability $p(S \text{ is culprit} | \text{ID filler}) = .32$. Maximum information gain for filler identifications with the mismatch-description method occurred when the prior probability $p(S \text{ is culprit}) = .57$ (at which point information gain was .13). Greater information gain for identifications of fillers for the match-description filler-selection method means the chance that the suspect was innocent when a filler was identified was greater when fillers were selected using the match-description method than when the mismatch-description method was used.

Information gain for "don't know" responses. Information gain for "don't know" responses was calculated for both filler-selection methods. For the match-description method, the average rate of "don't know" responses was .27 when the suspect was not the culprit and .15 when the suspect was the culprit. For the mismatch-description method, the average rate of "don't know" responses was .15 when the suspect was not the culprit and .07 when the suspect was the culprit. Information-gain curves for "don't know" responses for both of these filler-selection methods are shown in Figure 3D.

Notice that information gain for "don't know" responses was well above zero and was in the direction of exonerating evidence. This means that eyewitnesses were more likely to say "don't know" when the lineup did not contain the culprit than when the lineup contained the culprit. We know of no other work that has examined the possibility that "don't know" responses have any informational value. We note as well that the two filler-selection methods appear to have yielded similar amounts of information gain from "don't know" responses. For each method, eyewitnesses were approximately two times more likely to say "don't know" if

they observed a lineup that did not contain the culprit than if they observed a lineup that did contain the culprit.

Maximum information gain for "don't know" responses occurred at similar prior probability points for the two filler selection methods (from .57 to .62) and yielded similar peak levels of information gain (from .15 to .24).

Summary. The information gain advantage of the match-description filler-selection method over the mismatch-description filler-selection method surfaced primarily in identifications of the suspect and filler identifications. The mismatch-description filler-selection method performed rather poorly in terms of information gain regardless of the response of the eyewitness.

Comparisons of the Informational Values of Eyewitness Behaviors

In the previous two sections, we compared the informational value of one type of lineup with another as a function of the behavior of the eyewitness. In this section, we compare the informational value of the eyewitness behaviors with each other within lineup type. Our analyses comparing the informational value of different eyewitness behaviors uses only the simultaneous lineup data in which fillers were selected using the match-description method and the sequential lineup data. There are two reasons we used only these two types of lineups for comparing the informational value of different eyewitness behaviors. First, these two types of lineups are the only two recommended by eyewitness researchers and adopted in the recent national guidelines (Technical Working Group for Eyewitness Evidence, 1999). Hence, we expect the patterns that emerge with these two types of lineups to reflect patterns emerging from recommended lineup procedures. Second, only these two types of lineups showed clear separation of information gain between the various behaviors. The greater separation of the information-gain curves across eyewitness behaviors for the better types of lineups is what would be expected. When poor procedures (such as the mismatch-description filler-selection method) are used, the informational value of all eyewitness behaviors regresses toward zero, diminishing the differences in informational value that these behaviors carry. In general, we believe the information-gain pattern across behaviors is best studied with the best lineup types. Figure 4A shows the information-gain curves for the three behaviors of witnesses using the sequential lineup data. Here, we see that information gain for identifications of the suspect was greater than that for identifications of a filler, which in turn was greater than that for "not there" responses. The low informational value of "not there" responses relative to filler identifications is particularly interesting. Intuitively, we might expect that explicit "not there" responses would more clearly indicate that the culprit is not in the lineup than would filler identifications. Recall, however, that "don't know" responses are not collected with the sequential lineup. In large part, this is because the procedure involves sequential decisions to identify or not identify a given lineup member without knowing how many lineup members there will be. Using the sequential procedure, there is no point at which the witness is given an opportunity to say "don't know." We could speculate, therefore, that some of the failures to identify anyone by the end of the sequential procedure, which we have classified as "not there" responses, could be cases of "don't know" witnesses. Because we would expect "don't

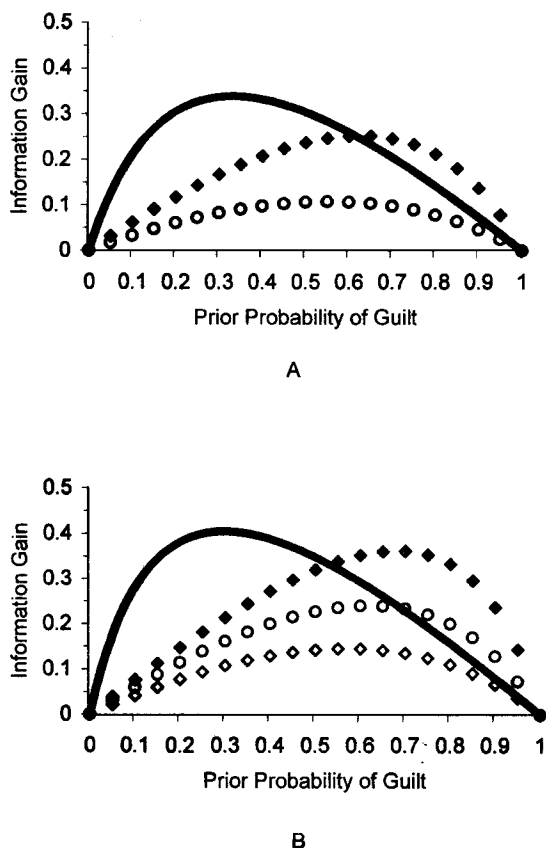


Figure 4. A: Information gain from each eyewitness behavior for the sequential lineup. B: Information gain from each eyewitness behavior for the match-description filler-selection strategy. In Figure 4A, solid lines = ID of suspect; solid squares = ID of filler; open circles = no ID. In Figure 4B, solid lines = ID of suspect; solid squares = ID of filler; open circles = "not there" responses; open squares = "don't know" responses.

know" responses to carry less informational value than "not there" responses, their inclusion in the "not there" category for the sequential lineup probably dilutes the informational value of "not there" responses.

Unlike the sequential lineup data, the match-description filler-selection method data partition the eyewitness responses into four behaviors: identification of the suspect, identification of a filler, "not there," and "don't know." The informational value of these four behaviors is shown in Figure 4B. As Figure 4B shows, even when "don't know" eyewitnesses are partitioned out of the "not there" response category, "not there" responses were less informative than were filler identifications. We believe that the rank order of information gain for the four behaviors is somewhat surprising because intuition suggests that explicit "not there" responses would be more informative of the innocence of the suspect than would filler identifications. We believe that this has important implications.

The surprising informational value of filler identifications. Perhaps the most provocative finding in these analyses is that filler identifications were more informative than the suspect was innocent than were explicit "not there" responses. Because this finding is counterintuitive at first glance, it deserves some discussion.

First, we should note that some studies show filler identifications to be more informative than "not there" responses (e.g., Juslin et al., 1996), whereas other studies do not show this pattern (e.g., Tunnicliff & Clark, 2000). Hence, although we do not know what governs the relative information gain of filler identifications versus "not there" responses, we do not believe it is always the case that filler identifications are more informative than "not there" responses. However, we think the exonerating informational value of filler identifications makes good sense. In effect, a filler identification is an eyewitness's way of saying that there is a filler who looks more like the culprit than does the suspect, which, of course, should be more likely to happen when the lineup does not contain the culprit than when the lineup does contain the culprit.

Once again, we argue information-gain analyses can reveal important patterns that otherwise tend to go unnoticed. For instance, the potential implications of our observation about the informativeness of filler identifications are enormous. Consider the likely reaction of a police investigator as a function of whether the witness made a "not there" response versus a filler identification. If the eyewitness viewed the lineup and responded "not there," the investigator would be likely to reduce his or her subjective probability that the suspect is the actual culprit. Suppose, however, the eyewitness selected a filler. A likely reaction by the investigator to the witness's behavior would be to decide that this is an unreliable witness and exhibit little or no change in his or her beliefs about the guilt of the suspect. At one level, the investigator is clearly justified in this reaction; after all, the witness has clearly made an error. However, as noted above, the error of identifying a filler is actually informative about the likelihood that the suspect is the culprit because a filler identification indicates there is someone else in the lineup (a filler) who resembles the culprit more than does the suspect. Simply passing off the witness who makes a filler identification as unreliable without also reducing significantly one's subjective probability that the suspect is the culprit would be a serious misreading of the information. Wright and McDaid (1996) analyzed the eyewitness identification patterns for 1,561 actual eyewitnesses and found that fillers were picked about 20% of the time. Hence, any tendency to dismiss filler identifications as uninformative represents a loss of useful exonerating information in a relatively high percentage of cases.

Of interest, although there appears to be rather good agreement among eyewitness experts regarding variables that affect eyewitness identification accuracy (Kassin, Tubb, Hosch, & Memon, 2001), there has been virtually no dialogue in the eyewitness identification literature regarding the meaning of exonerating behaviors ("not there" and filler identification responses). Use of the information-gain approach to eyewitness identification data can help correct this oversight and should result in eyewitness identification researchers' routinely examining the informational value of exonerating behaviors.

Information-gain curve intersections. The information-gain curves in Figures 4A and 4B permit us to make another important observation. Because incriminating behaviors (identification of suspect) and exonerating behaviors ("not there," filler identifications, and "don't know") have information-gain curves that skew in opposite directions, they always intersect. These intersections represent points on the prior probability variable at which the two forms of information are equally informative: Below that intersecting point, the incriminating behavior is more informative than

is the exonerating behavior, but above that intersecting point, the exonerating behavior is more informative than is the incriminating behavior. For instance, notice in Figure 4B that the information-gain curve for identification of suspect intersected the information-gain curve for filler identifications when the prior probability $p(S$ is culprit) was .54. This means that an identification of a filler was actually more informative than was the identification of a suspect when the prior probability $p(S$ is culprit) was greater than .54. Similarly, we can determine from the data in Figure 4B that the informational value of a "not there" response exceeded the informational value of an identification of the suspect when the prior probability $p(S$ is culprit) exceeded .70.

Because information-gain curves for incriminating behaviors always intersect with information-gain curves for exonerating behaviors, we can never conclude that an incriminating behavior (identification of suspect) is more informative than an exonerating behavior (e.g., "not there") or vice versa unless we also make reference to a specific prior probability, prior probability region, or some type of average over the range of possible priors. This observation reinforces two of our main points about information-gain analyses: The information-gain approach to eyewitness identification data helps flesh out assumptions and helps reveal qualifications to conclusions that would not be apparent from traditional analyses.

Significance Testing

How does a researcher determine whether the area under an information-gain curve is significantly different from zero or whether one information-gain curve is significantly different from another information-gain curve? The answer is actually quite simple as long as researchers keep in mind that significance tests are based solely on the proportions derived from the cell-partitioned frequencies in the experiment, not the information-gain curves. Consider, for instance, the information-gain curve for the sequential lineup in Figure 2A. That curve was derived from the proportion of witnesses who selected the suspect when the suspect was the culprit (47% of the 113 in this condition who viewed a culprit-present lineup) and the proportion who selected the suspect when the suspect was not the culprit (12% of the 113 in this condition who viewed the culprit-absent lineup). A test of the differences in these proportions is equivalent to the test of whether information gain from identifications of the suspect for the sequential lineup is different from zero. If we use the arcsine transformation (ϕ) method (Winer, 1962), a simple test of this is provided by the ratio

$$z = \frac{(\phi_a - \phi_b)}{\sqrt{1/N + 1/N}}$$

In the case of the information-gain curve for identification of suspect with the sequential lineup,

$$z = \frac{(1.51 - .708)}{.133} = 6.04.$$

The z value of 6.04 is significant at $p < .01$, indicating that the curve is significantly greater than zero using conventional tests of statistical significance.

A simple extension of the above logic permits tests of the difference between two or more information-gain curves. The question of whether one information-gain curve is different from another information-gain curve is equivalent to an interaction test on the four proportions involved. Consider, for instance, the difference between the simultaneous and sequential curves in Figure 2A. These curves were derived from the fact that the suspect was identified by 55% of the witnesses in the simultaneous culprit-present lineup condition, 32% of the witnesses in the simultaneous culprit-absent lineup condition, 12% of the witnesses in the sequential culprit-absent lineup condition, and 47% of the witnesses in the sequential culprit-present lineup condition. There were 113 witnesses in each of these four conditions. The interaction test based on the arcsine transformation method is simply

$$z = \frac{(\phi_a - \phi_b) - (\phi_c - \phi_d)}{\sqrt{1/N + 1/N + 1/N + 1/N}}$$

In this case, the interaction is significant using a one-tailed test, $z = 1.76$, $p < .05$, supporting the prediction that information gain for identifications of the suspect was higher for the sequential than for the simultaneous lineup procedure.

Effect sizes can be calculated in the traditional manner from the z scores. However, it should be noted that the effect size associated with the z score assumes the base rate for culprit-present and culprit-absent lineups is 50%, as it was in the experiment itself. Actual effect sizes in the real-world situation would depend on the base rate. This is apparent from considering the extremes of the base-rate variable. For instance, if the base rate for culprit-present lineups were 100%, then the effect size for the difference in information gain between simultaneous and sequential lineups is zero because neither lineup would have information-gain value.

Coda

We have outlined an approach to the analysis of eyewitness identification data that we strongly encourage eyewitness identification researchers to consider. Like many applications of mathematical models to data, information-gain analyses force a systematic articulation and examination of hidden assumptions and impose order on various representations of the problem. The information-gain analysis we used here, for example, naturally elicited a full consideration of all possible eyewitness behaviors in response to a lineup (in part because the sum of the probabilities of these behaviors must add up to 1.00). We are not suggesting that all written reports of eyewitness identification experiments must or even should include information-gain curves. We believe, however, that researchers would find it useful to examine their data in ways that consider the impact of prior probabilities or base rates. The analyses of previously published experimental data we reported here, for instance, revealed some provocative patterns that were not noticed in the original publications and could not have been discovered with traditional analyses. Particularly interesting was the way the prior probability variable interacts with witness behavior in terms of information gain. Even when a behavior appears to have only modest exonerating value, it could have more informational value than does an incriminating behavior if the base rate for the culprit's being in the lineup is high.

Another advantage of the information-gain approach to eyewitness identification data is that it keeps salient that there is an

important parameter (the base rate for culprit-absent lineups) that is not represented in experiments and that we know little about in the real world. This should give researchers pause and introduce some conservatism in generalizing to the real world. For instance, it is tempting to conclude from experimental data that, compared with the simultaneous lineup, the sequential lineup reduces the chances by 50% that the identification of a suspect is a mistaken identification. However, the actual percentage depends strongly on the base rate for culprit-present and culprit-absent lineups.

The information-gain approach also helps encourage eyewitness researchers to learn more about the factors controlling the culprit-present versus culprit-absent base rate in the real world and perhaps motivate researchers to find plausible upper and lower bounds of this base rate. Along this line, we offer some initial thoughts.

First, we note it is commonly true in real cases that a given eyewitness would view more than one lineup even though there was only one culprit. In these cases, we can assume the base rate for the actual culprit's being in a given lineup can be no greater than $1/N$ where N is the number of lineups viewed by that eyewitness. Hence, one approach to setting an upper limit on the base rate for culprit-present lineups is to collect real-world data on how many single-suspect lineups the average eyewitness is shown relative to the number of culprits who these suspects are hypothesized to be. If that ratio was 2:1, for instance, then the base rate probability that a lineup's suspect is the culprit could not be greater than .50. The probability could be less than .50, of course, but this figure would at least establish an upper boundary.

Second, researchers could try to study the factors that affect the base rate for culprit-present versus culprit-absent lineups in the real world. Such research would likely reveal considerable variation from one police department to another. In some departments, for instance, police might require some rather strong reasons to suspect a person before they would place that person in a lineup. In other departments, police might place a person in a lineup based on nothing more than a hunch. Over time, these two departments could end up with very different base rates for culprit-present versus culprit-absent lineups and correspondingly different amounts and patterns of information gain. Indeed, we believe the base rate for culprit-present and culprit-absent lineups in the real world should not be construed as a single "true" number. Instead, this base rate is a variable that will change as a function of the prelineup investigative practices of police and the criteria they use for deciding when to consider someone a plausible enough suspect to conduct a lineup.

The legal system is now paying a great deal of attention to experimental eyewitness identification research in psychology (Wells et al., 2000). Much of this close attention is the result of recent DNA exoneration cases, which show that mistaken eyewitness identification was the primary cause of the convictions of these innocent people (Connors, Lundregan, Miller, & McEwan, 1996; Scheck et al., 2000; Wells et al., 1998). We believe that a consideration of the prior probability variable is a necessary next step in further advancing this interface with the legal system. Information-gain analyses permit us to examine the role of the prior probability variable without having to know what the prior probability is in the real world. This strikes us as far preferable to being base-rate neglectful.

References

- 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* (NCJ No. 161258). Alexandria, VA: National Institute of Justice.
- Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73*, 281-290.
- Dawes, R. M. (1996). The purpose of experiments: Ecological validity versus comparing hypotheses. *Behavioral and Brain Sciences, 19*, 20-21.
- Doob, A. N., & Kirshenbaum, H. (1973). Bias in police lineups—Partial remembering. *Journal of Police Science and Administration, 1*, 287-293.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249-267). New York: Cambridge University Press.
- Goldstein, A. G., Chance, J. E., & Schneller, G. R. (1989). Frequency of eyewitness identification in criminal cases: A survey of prosecutors. *Bulletin of the Psychonomic Society, 27*, 71-74.
- Justin, P., Olson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can and cannot be inferred from a low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 5*, 1304-1316.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237-251.
- Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research. *American Psychologist, 56*, 405-416.
- Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*, 1-53.
- Levi, A. M. (1998). Are defendants guilty if they were chosen in a lineup? *Law and Human Behavior, 22*, 389-408.
- Lindsay, R. C. L., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76*, 741-745.
- Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection method. *Law and Human Behavior, 18*, 527-541.
- Lindsay, R. C. L., & Wells, G. L. (1980). What price justice? Exploring the relationship between lineup fairness and identification accuracy. *Law and Human Behavior, 4*, 303-314.
- Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentations. *Journal of Applied Psychology, 70*, 556-564.
- Luus, C. A. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15*, 43-57.
- Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482-489.
- Melara, R. D., DeWitt-Rickards, T. S., & O'Brien, T. P. (1989). Enhancing lineup identification accuracy: Two codes are better than one. *Journal of Applied Psychology, 74*, 706-713.
- Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence*. New York: Random House.
- Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78*, 22-33.
- Swets, J. A., Tanner, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review, 68*, 301-340.

- Technical Working Group For Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* (NCJ No. 178240). Washington, DC: U.S. Department of Justice, Office of Justice Programs.
- Tunnick, J. L., & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions? *Law and Human Behavior, 24*, 231-258.
- Wells, G. L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*, 1546-1557.
- Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48*, 553-571.
- Wells, G. L., Leippe, M. R., & Ostrom, T. M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior, 3*, 285-293.
- Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88*, 776-784.
- Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55*, 581-598.
- Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). On the selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*, 835-844.
- 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.
- Wells, G. L., & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99*, 320-329.
- Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- Wright, D. B., & McDaid, A. T. (1996). Comparing system and estimator variables using data from real lineups. *Applied Cognitive Psychology, 10*, 75-84.

Received January 10, 2002

Revision received April 8, 2002

Accepted April 9, 2002 ■

New Editors Appointed, 2004-2009

The Publications and Communications Board of the American Psychological Association announces the appointment of five new editors for 6-year terms beginning in 2004.

As of January 1, 2003, manuscripts should be directed to the following individuals. Authors are strongly encouraged to submit manuscripts electronically as of January 1, 2003, through the journal's Manuscript Submission Portal (see the Web site listed with each journal title).

- For *Psychology and Aging* (<http://www.apa.org/journals/pag.html>), submit manuscripts to **Rose T. Zacks, PhD**, Department of Psychology, Michigan State University, East Lansing, MI 48824-1117.
- For *Psychological Assessment* (<http://www.apa.org/journals/pas.html>), submit manuscripts to **Milton E. Strauss, PhD**, Department of Psychology, Case Western Reserve University, Cleveland, OH 44106-7123.
- For *Journal of Family Psychology* (<http://www.apa.org/journals/fam.html>), submit manuscripts to **Anne Kazak, PhD, ABPP**, Oncology Psychosocial Research, The Children's Hospital of Philadelphia, Room 1486 (Market Street), 34th and Civic Center Boulevard, Philadelphia, PA 19104. For overnight couriers: Room 1486, 3535 Market Street, Philadelphia, PA 19104.
- For *Journal of Experimental Psychology: Animal Behavior Processes* (<http://www.apa.org/journals/xan.html>), submit manuscripts to **Nicholas Mackintosh**, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, United Kingdom.
- For *Journal of Personality and Social Psychology: Personality Processes and Individual Differences* section (<http://www.apa.org/journals/psp.html>), submit manuscripts to **Charles S. Carver, PhD**, Department of Psychology, University of Miami, P.O. Box 248185, Coral Gables, FL 33124-2070.

Manuscript submission patterns make the precise date of completion of the 2003 volumes uncertain. Current editors Leah L. Light, PhD, Stephen N. Haynes, PhD, Ross D. Parke, PhD, Mark E. Bouton, PhD, and Ed Diener, PhD, respectively, will receive and consider manuscripts through December 31, 2002. Should 2003 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2004 volumes.