HANDBOOK OF EYEWITNESS PSYCHOLOGY

Volume 2

MEMORY FOR PEOPLE

EDITED BY

Rod C. L. Lindsay
Queen's University, Ontario

David F. Ross
University of Tennessee, Chattanooga

J. Don Read
Simon Fraser University, British Columbia

Michael P. Toglia
State University of New York, Cortland

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
2006
Mahwah, New Jersey London
The notion of a theory of lineups is debatable and, in the minds of many psychological theorists, might seem peculiar. Psychological theories usually concern circumscribed psychological processes. For instance, there are theories of face recognition, theories of judgment and decision making, theories of attitude formation and change, and so on. In contrast, a lineup is a task or a procedure, and, although it involves psychological processes, it is not a psychological process per se. Nevertheless, we contend that it is possible to describe some necessary components of a theory of lineups. In particular, we propose various considerations that ought to be a part of an applied theory of lineups. We call it an applied theory because it is embedded in the ecology of the legal system and because it rests heavily on the applied goals of the users of lineups, principally the goal of incriminating the guilty and exculpating the innocent.

What qualities should an applied theory of lineups have? We approached this problem with the notion that an applied theory of lineups should describe the function of lineups, the structural properties of lineups and their resultant outcome distributions, and the broad psychological processes that are presumed to operate during a lineup task. In addition, in order to be useful, an applied theory ought to help illuminate gaps in our empirical knowledge, predict patterns of data not yet observed, and stimulate new research.

We thoroughly ground our lineup theory in an ecology that mimics the constraints of the real-world of police investigation and speaks to the question of courtroom probative value. These constraints have significant implications for any applied theory. Consider an example. Traditional theoretical approaches to recognition memory in psychology focus on estimating the probability of observing a particular response (e.g., positive identification) given the presentation of a particular stimulus (e.g., a previously seen stimulus or a novel stimulus). Real-world models of the lineup, on the other hand, need to inform us about the probability that a particular stimulus was presented given that a particular response was observed. As discussed in the section “Bayesian Distribution of Lineup Outcomes,” this distinction between the probability of observing a response given the presence of a stimulus versus the probability that a stimulus is present given the observation of a response is critical to a practical theory of lineups. Consider another
difference between traditional theoretical approaches to memory and an applied theory of lineups. In traditional psychological theorizing regarding recognition memory, the mistaken identification of a novel stimulus from a display that included the original (correct) stimulus is not fundamentally different from the identification of a novel stimulus from a display that did not include the original (correct) stimulus. As explained later in this chapter, however, there are very different types of errors in lineups that belong in different outcome categories because they can have vastly different consequences. Also, unlike traditional theorizing about recognition memory, a lineup theory must include a role for social influence and for inference processes that are not, in and of themselves, memory variables.

We begin our attempt to articulate a theory of lineups by describing the function of lineups. Why is a lineup conducted in actual cases? What is the goal? For example, is the goal to test the memory of the eyewitness? We contend that the answer is no and that the goal instead is to test a hypothesis regarding whether the suspect is the actual perpetrator. We then describe the structure of lineups. Although there are various possible structures, some elements are a necessary part of the structure, such as the perpetrator-present versus perpetrator-absent element. After establishing the function and structure of lineups, outcome categories can be meaningfully described. Here, for example, we distinguish between known errors and unknown errors. We then describe broad psychological processes that can operate to affect the ability of the lineup to serve its intended function of testing the hypothesis that the suspect is the perpetrator.

THE FUNCTION OF LINEUPS

If you asked 10 people why police conduct lineups, most would say something like “To see if the eyewitness can recognize the bad guy” or something along that line. Indeed, many eyewitness identification researchers are saying something similar when they describe a lineup as a test of the witness’s memory. We could not disagree more. Although it can be the purpose of an experiment to test the ability of eyewitnesses to identify targets under particular conditions, this is not the purpose of a lineup in a criminal investigation. Instead, the purpose of a lineup in a criminal investigation is to gather evidence regarding whether the suspect in the lineup is the actual perpetrator.

The distinction between these two functions of a lineup, one testing the witness and the other testing a hypothesis regarding the suspect, is not mere semantics. Although crime investigators can and do make inferences about the reliability of an eyewitness as a function of the eyewitness’s response to a lineup, their main objective is to learn something about the likely guilt of the suspect. So, if the eyewitness picks their suspect from the lineup, investigators may increase their certainty that the suspect is in fact the culprit in question. On the other hand, if the witness does not identify the suspect from the lineup, investigators may decrease their certainty that the suspect is the perpetrator. In the latter case, crime investigators could then broaden the investigation to consider other possible suspects. Whether investigators use the information from the witness appropriately is not the point. For example, it is quite possible that criminal investigators
too readily dismiss nonidentifications for their exonerating qualities while readily accepting identifications of the suspect for their incriminating qualities (see Wells & Olson, 2002). But this does not negate the point that the purpose of the lineup was to test the hypothesis that the suspect is the perpetrator.

If the actual purpose of a lineup in real cases is to test the memory of the witness, then we would expect to see some things happening that are, on their face, quite absurd. For example, we would expect to have the greatest interest in conducting a lineup when there already exists definitive proof beyond all doubt that the suspect is the criminal. So, if there were a confession from the suspect, plus definitive DNA evidence and fingerprints, investigators would definitely want to conduct a lineup, because that would be the best situation for testing the memory abilities of the eyewitness. After all, that situation would be much more informative about the abilities of the eyewitness than would a situation in which there was uncertainty about whether the suspect is the perpetrator. Notice, however, that it does not make much sense to bother with a lineup under these circumstances, because the actual purpose of a lineup is not to test the memory of the witness but rather to test the hypothesis that the suspect is the perpetrator. Likewise, suppose that the case was seemingly unsolvable because there was no suspect whatsoever. Again, if the purpose of a lineup is to test the eyewitness’s memory abilities, investigators would want to use a blank lineup, which is a lineup in which all members are known innocents, to see whether the eyewitness will choose someone nevertheless. After all, a lineup with no suspect would serve as a test to tell us something about whether the eyewitness is reliable. Of course, even though the memory abilities of the eyewitness are clearly relevant at some level, it is not the purpose of the lineup to test the memory abilities of the eyewitness. These examples are absurd precisely because they miss the real function of a lineup, which is to help determine whether the suspect is the perpetrator.

The test-of-witness versus test-of-suspect-guilt distinction is more important than it might first appear, as explained later in this section. At the same time, it is quite understandable that psychological researchers have commonly thought of the function of a lineup as being a test of the eyewitness. Psychological research is almost always a test of how a subject will respond to a stimulus of one type versus another type rather than a test of whether a stimulus is of one type versus another type based on the subject’s response. In a signal detection task, for example, a researcher is interested in what response a subject makes to a stimulus condition (noise alone versus signal plus noise) where the stimulus condition is known with certainty. A signal detection researcher is never in the position of trying to decide what the stimulus condition was (noise alone versus signal plus noise) based on the response of the witness. In this sense, the forensic nature of the lineup task differs significantly from the nature of a signal detection experiment because the “givens” and the “unknowns” are different in criminal investigations from what they are in an eyewitness experiment. An eyewitness experiment, for example, is designed to assess the probability of a particular response from the eyewitness, given that the witness was presented with one stimulus (the actual culprit) versus a different stimulus (an innocent suspect) and the experimenter knows the status of the stimulus (i.e., whether the stimulus is the culprit or not). A real-world lineup case, in contrast, is designed to assess the probability that one stimulus (the actual culprit) versus a different stimulus (an inno-
cent suspect) was presented to the eyewitness, given a particular response of the witness. In the language of conditional probabilities, the two situations rely on opposite conditionals: Experiments are concerned with the probability of R given S, whereas real-world eyewitness cases are concerned with the probability of S given R (where R and S mean response and stimulus, respectively).

What difference does it make to have one construal of the conditional or the other? It can make a considerable difference, and we offer some numerical examples to illustrate our point. Suppose that we conducted an experiment in which we asked people to indicate whether or not a face shown to them was a former president of the United States. We use 20 faces of former presidents and 10 of nonpresidents, for a total of 30 faces. Suppose that we obtained results like those shown in Table 10.1. As indicated in Table 10.1, of the 20 times a president’s face was shown, the subject said it was a president 14 times (70% hit rate), and of the 10 times a nonpresident’s face was shown, the witness said it was not a president 7 times (70% correct rejections). So, the probability of a correct response given a particular condition (i.e., saying “president” to a president’s face, or saying “not president” to a nonpresident’s face) is 70%. But what is the probability that a stimulus was of a particular type (president versus nonpresident), given a particular response? For example, what is the probability that the subject was looking at a president’s face, given that the subject said “president”? There were 17 times that a subject said “president,” during 14 of which the picture shown was of a president. Hence, the correct answer is 82.4% (14/17). And, what is the probability that a subject was looking at a nonpresident, given that the subject said “not a president”? There were 13 times in which the subject said “not president,” 7 of which the picture shown was not of a president. Hence, the correct answer is 53.8% (7/13). Clearly, the probability of a response given a stimulus is not the same as the probability of a stimulus given a response. Criminal investigators and triers of fact in criminal cases are in the latter position: estimating the probability that the suspect is guilty or innocent given the response of the witness, not the probability of the witness’s response given that the suspect is innocent or guilty. This directional distinction is not arbitrary but rather is necessary, given the forensic ecology under which real-world lineups are conducted.

It is important to note that our depiction of the function of lineups (to test a hypothesis regarding the innocence or guilt of the suspect) is bidirectional. When properly designed and interpreted, a lineup procedure has not only incriminating powers, but exonerating powers as well. In fact, there is clear proof using mathematical formulations that any lineup that has incriminating value from the identification of the suspect must also have exonerating value from a nonidentification (Wells & Lindsay, 1980; Wells & Olson, 2002). The magnitude of the incriminating value does not have to be mathematically

<table>
<thead>
<tr>
<th></th>
<th>President</th>
<th>Not President</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject says “a president”</td>
<td>14</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Subject says “not a president”</td>
<td>6</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>10</td>
<td>30</td>
</tr>
</tbody>
</table>
equal to the magnitude of the exonerating value; one can be greater than the other, depending on several factors that we discuss later in this chapter.

THE STRUCTURE AND OUTCOMES OF LINEUPS

Lineup Structure

In the most general terms, a lineup is structured so that a suspect is embedded among nonsuspects. We will call these nonsuspects fillers. In some eyewitness identification writings, fillers are known as foils or distractors. In an episode of Seinfeld, fillers were called decoys. We prefer the term fillers because this appears to be the most common term used by law enforcement in the United States, and it seems rather neutral (foils, for instance, seems to imply that their purpose is to fool the witness). Whatever they are called, the important thing about fillers is that they are not suspects, but rather are known innocents. Accordingly, any identification of a filler by an eyewitness is a known error that would not result in charges against that person.

It is critical to maintain a distinction between a suspect and a perpetrator. A suspect is someone who is suspected of being the perpetrator, but there are innocent suspects and there are guilty suspects. Indeed, as discussed in the previous section, the function of a lineup is to bring evidence to bear on the question of whether the suspect and the perpetrator are the same person or are different people. Experiments on eyewitness identification represent these two types of suspects (the guilty suspect and the innocent suspect) by showing some eyewitnesses lineups that include the perpetrator and showing other eyewitnesses lineups that do not include the perpetrator. Perpetrator-present and perpetrator-absent lineups are a necessary structural feature of the design of any eyewitness identification experiment concerned with eyewitness identification accuracy precisely because they are part of the ecology of real-world lineups.

Outcomes Taxonomy

We think it is useful to depict a lineup as having eight possible outcomes. These outcomes derive from a combination of two states of truth (perpetrator present versus perpetrator absent) crossed with four possible behaviors of the eyewitness in response to the lineup. The four behaviors are the identification of the suspect, the identification of a filler, a “not present” response, and a “don’t know” response.

1Not all law enforcement agencies have followed this fundamental structure. In some cases, lineups have been composed entirely of suspects and no member of the lineup was a filler, even though there was only one perpetrator (Wells & Turtle, 1986). Nevertheless, our theory assumes a lineup is structured to include only one suspect and the remaining members are fillers, which is consistent with guidelines from the U.S. Department of Justice (Technical Working Group on Eyewitness Evidence, 1999).

2There are, of course, other possible responses that might make outcome classification difficult. For example, an eyewitness might say, “I think it is number three, but I’m not sure.” Would this be an identification of number three or would this be a “don’t know” response? Alternatively, an eyewitness might say, “It might be number four, but it could be number six.” Our taxonomy does not itself tell us how to treat such responses.
Table 10.2 displays the eight possible outcomes. In actual cases, only the response of the witness is known with certainty, whereas in experiments the state of truth (perpetrator-present or perpetrator-absent) is also known. A miss occurs when the perpetrator is present but the eyewitness fails to identify the perpetrator. Notice that we have depicted three types of misses. A Type 1 miss occurs when the perpetrator is in the lineup and the eyewitness says that the perpetrator is "not present." A Type 2 miss occurs when the perpetrator is present and the eyewitness identifies a filler instead of the perpetrator. A Type 3 miss occurs when the perpetrator is present and the eyewitness says "don't know." A correct rejection occurs when the perpetrator is not in the lineup and the eyewitness does not identify the innocent suspect. Notice that there are three types of correct rejections. Correct Rejection Type 1 occurs when the perpetrator is absent and the eyewitness says "not present." A Correct Rejection Type 2 occurs when the perpetrator is absent and the witness identifies a filler. A Correct Rejection Type 3 occurs when the perpetrator is absent and the eyewitness says "don't know."

Undoubtedly, some readers would disagree with our characterization of Correct Rejection Types 2 and 3 as being rejections of anything. In the case of Type 2 Correct Rejection, the eyewitness clearly does not reject the lineup, but instead identifies a known-innocent filler. Nevertheless, with a Type 2 Correct Rejection, the eyewitness has correctly rejected the idea that the suspect is the perpetrator and it is in this sense that a Type 2 Correct Rejection is both correct and a rejection. Likewise, readers might find it odd to call a "don't know" response to a perpetrator-absent lineup a Correct Rejection Type 3. However, even though the eyewitness did not explicitly say that the perpetrator was not present, the eyewitness nevertheless made a correct decision to not identify the suspect.

Outcome Distributions

Outcome distributions are simply the frequencies, percentages, or probabilities associated with the eight outcome cells depicted in Table 10.2. This is the meat of the matter of eyewitness identification accuracy, because it is these distributions that define the ability of the lineup task to inform us about the hypothesis that the suspect is the perpetrator.

Note that we can partition the columns in various ways by collapsing some columns and not others. Particularly useful is collapsing the "not present," identification of filler, and "don't know" columns into a single column that we could label "non-identifications.

<table>
<thead>
<tr>
<th>Perpetrator-present lineup</th>
<th>Identification of Suspect</th>
<th>&quot;Not Present&quot; Response</th>
<th>Identification of Filler</th>
<th>&quot;Don't Know&quot; Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perpetrator-absent lineup</td>
<td>Accurate identification</td>
<td>False rejection or Type 1 miss</td>
<td>Type 2 miss</td>
<td>Type 3 miss</td>
</tr>
<tr>
<td></td>
<td>Mistaken identification</td>
<td>Correct rejection Type 1</td>
<td>Correct rejection Type 1</td>
<td>Correct rejection Type 3</td>
</tr>
</tbody>
</table>

TABLE 10.2
of the suspect.” For economy of language, we will simply call these nonidentifications, even though they include identifications of fillers. Collapsing the nonidentifications together is a useful tool for applying some simple mathematics because it leaves only two categories of response, identifications of the suspect and nonidentifications.

Consider, for example, the hypothetical distribution of responses in Table 10.3. Readers can think of these numbers as frequencies that resulted from showing a perpetrator-present lineup to 100 witnesses and a perpetrator-absent lineup to 100 witnesses in which 60 accurately identified the suspect from the perpetrator-present lineup, 40 did not, and 20 mistakenly identified the innocent suspect from the perpetrator-absent lineup, whereas 80 did not. Alternatively, readers can think of these as probabilities in which the probability was .60 that the witness would make an accurate identification of the suspect from the perpetrator-present lineup and the probability was .40 that they would not, and so on. Or, these could be construed as percentages in which 60% accurately identified the suspect from the perpetrator-present lineup, whereas 40% did not and so on. In any case, these numbers can be used to form a ratio within types of response (i.e., within the columns). Now, suppose we sample a witness at random and, as in an actual case, we do not know whether the suspect was the perpetrator or not. Suppose, instead, that all we know is that the witness identified the suspect. Given an identification of the suspect, what is the probability that the suspect is in fact the perpetrator? The answer is 75%, because 60 of the 80 who identified a suspect identified the perpetrator (i.e., were viewing a perpetrator-present lineup). What if all we knew was that the witness made a nonidentification? What is the probability now that the suspect is the perpetrator? The answer is 33.3%, because 40 of the 120 witnesses who made nonidentifications were viewing a perpetrator-absent lineup.

A somewhat preferred way of expressing the information in Table 10.3 is in the form of diagnosticity ratios. The diagnosticity ratio for a response is an indication of how much more likely one thing is than another when that response occurs. For instance, note that the identification of the suspect is three (3.0) times more likely when the suspect is the perpetrator than it is when the suspect is not the perpetrator. Likewise, the diagnosticity of a nonidentification in Table 10.3 is two (2.0) because it is twice as likely that a nonidentification response will occur when the suspect is not the perpetrator than when the suspect is the perpetrator. A diagnosticity ratio of 1.0 indicates that there is no diagnosticity at all. The higher the diagnosticity ratio, the more impact the response of the witness should have on the hypothesis that the suspect is the perpetrator.

This leads us to our first, and undoubtedly least controversial, observation about the distribution of outcomes as it relates to diagnosticity: Any lineup that has incriminating diagnosticity resulting from an identification of the suspect must also have exonerating diagnosticity resulting from an identification of the suspect.

<table>
<thead>
<tr>
<th></th>
<th>Identification of Suspect</th>
<th>Non-Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perpetrator-present lineup</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>Perpetrator-absent lineup</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>
diagnosticity from a nonidentification. The reason that this observation ought not to be controversial is that it is a mathematical necessity. The mathematical reason is simple. Because a witness must make either an identification or a nonidentification, the probability of a nonidentification has to be 1.0 minus the probability of an identification. Hence, whenever the probability of identification of a guilty suspect is greater than the probability of identification of an innocent suspect, the probability of nonidentification of an innocent suspect has to be greater than the probability of nonidentification of a guilty suspect.

The foregoing analysis tells us that nonidentifications have exonerating probative value, but it does not tell us how the exonerating value is distributed across the three types of nonidentification responses. One possibility is that exonerating diagnosticity is distributed equally across the three nonidentification responses. Another possibility is that one or two of the nonidentification response types have no exonerating value and the other one or two have exonerating value. There is nothing in the mathematics to constrain the possibilities, except that there must be exonerating value in at least one of the three nonidentification responses whenever there is incriminating value in an identification of the suspect. This is a domain in which only data collection can provide a defensible answer. Nevertheless, mere intuition and common sense would suggest that most, if not all, of the exonerating power of non-identifications resides in the “not there” response. This possibility is shown in Table 10.4. Notice in Table 10.4 that the diagnosticity ratio for “not there” responses is 3.0, but the diagnosticity ratios for filler identifications and “don’t know” responses is 1.0 (no diagnosticity). This is one domain, however, where mere intuition may be wrong. Analyses of data from eyewitness identification experiments tend to show outcome distributions more like those shown in Table 10.5 in which filler identifications have some diagnostic value and even “don’t know” responses have a small amount of diagnostic value (Wells & Lindsay, 1980; Wells & Olson, 2002).

**PERPETRATOR-PRESENT VERSUS ABSENT LINEUPS AS A PROBABILISTIC VARIABLE**

The perpetrator-present versus perpetrator-absent variable is a central and necessary aspect of any applied theory of lineups because it makes a profound difference to outcome distributions. In fact, under the assumption of a single-suspect lineup, it is impossible to have a mistaken identification of a suspect from a perpetrator-present lineup. The other reason that the perpetrator-present versus perpetrator-absent lineup variable is central and necessary to any applied theory of lineups is that it varies in actual cases.

**TABLE 10.4**

<table>
<thead>
<tr>
<th>Identification of Suspect</th>
<th>“Not Present” Identification of Filler</th>
<th>“Don’t Know” Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perpetrator-present lineup</td>
<td>60</td>
<td>10</td>
</tr>
<tr>
<td>Perpetrator-absent lineup</td>
<td>20</td>
<td>60</td>
</tr>
</tbody>
</table>
In this section, we show how this variable is perhaps best represented as a probabilistic variable, and we describe the kinds of factors likely to affect the present-absent variable in actual cases.

The present-absent variable is, at one level, quite straightforward in an experiment. Because the experimenter knows who committed the “crime,” it is a simple act to include this person in some lineups shown to witnesses and not include this person in other lineups shown to witnesses. In actual cases, it is not known whether the perpetrator is in the lineup or not. Indeed, the purpose of the lineup is to help make this determination based on how the witness behaves when confronted with the lineup. Hence, whereas the present-absent variable is an independent variable in an experiment, it is more like a dependent variable in actual cases. (This observation relates closely to the earlier distinction between assessing the likelihood of a witness making a particular response given the suspect is guilty, versus assessing the likelihood of guilt given that the witness makes a particular response.)

Like other dichotomous variables that are associated with uncertainty, the perpetrator-present versus perpetrator-absent variable can be construed as a probabilistic variable. In an experiment, it is common for the design to use a perpetrator-present lineup for half of the eyewitnesses and a perpetrator-absent lineup for the other half. Hence, for any given, randomly sampled lineup from the set of lineups in the experiment, the probability that the perpetrator is in the lineup (probability that the suspect is the perpetrator) is .50.

In our applied theory of lineups, in contrast, actual instances of lineups are not a random sample from some parent population of known proportions of perpetrator-present and perpetrator-absent lineups. There are several respects in which we reject the random sample from a known population idea of the present-absent variable in actual cases. First, the population of present versus absent lineups is unknown in actual cases. We acknowledge that methods might be developed to estimate the proportion of actual lineups in which the perpetrator was present versus absent. However, even if we had a large sample of past actual lineups from which we could somehow determine the proportion (base rates) of times that the perpetrator was or was not present, we would not consider this base rate to be stable across time or jurisdictions. This is due to the fact that the present-absent variable in actual cases is dynamically influenced by investigative practices and criteria that can vary dramatically over time and across jurisdictions.

Wells (1993) and Wells and Olson (2002) have described hypothetical law enforcement agencies that use lax versus strict criteria for deciding whether to put a suspect in a lineup. In a strict criterion jurisdiction, investigators might require that there be probable cause, such as motive, opportunity, tools of the crime, or some other fact that would

### Table 10.5

<table>
<thead>
<tr>
<th>Identification of Suspect</th>
<th>“Not Present” Response</th>
<th>Identification of Filler</th>
<th>“Don’t Know” Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perpetrator-present lineup</td>
<td>60</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Perpetrator-absent lineup</td>
<td>20</td>
<td>30</td>
<td>25</td>
</tr>
</tbody>
</table>
link the suspect to the crime before deciding to expose the witness to a lineup that includes that suspect. A lax criterion jurisdiction, on the other hand, might have no such requirement, and the presence of a mere hunch (or the inability to find other suspects) might be sufficient to decide to conduct a lineup with a particular suspect. Over time, the lax criterion and the strict criterion jurisdictions will run very different base rates for the present-absent variable. Because there is no legal requirement for probable cause to place a suspect in a lineup, especially with regard to the use of photographic lineups, the base rate for the present-absent factor can vary widely. Interestingly, there are times when it is possible to specify an upper limit on the present-absent base rate for a given investigation. For instance, suppose that there was one perpetrator and the eyewitness was shown two lineups, each with a different suspect. In this case, the maximum base rate across the two lineups is 50% for the witness viewing a perpetrator-present lineup. In other words, at least one of the two lineups had to be a perpetrator-absent lineup, so the 50% base rate is the upper limit.

Our applied theory of lineups cannot in itself specify a base rate for the present-absent variable, so the theory treats the present-absent factor as a continuous, probabilistic variable. Furthermore, because the present-absent factor is in fact a base rate, the present-absent factor combines with the diagnosticity factor (described in the previous section) in a manner described by Bayes' Theorem. Indeed, we are unable to imagine an applied theory of lineups that does not incorporate Bayesian parameters.

**BAYESIAN DISTRIBUTIONS OF LINEUP OUTCOMES**

Thus far we have laid a foundation for a Bayesian depiction of lineup outcomes. This foundation included (a) the distinction between estimating the probability of a witness response given that the suspect is or is not the perpetrator versus estimating the probability that the suspect is or is not the perpetrator given a response by the witness, (b) a description of the function of lineups as a test of the hypothesis that the suspect is the perpetrator rather than as a test of the witness's memory, (c) distinctions between various witness responses, and (d) the idea of the perpetrator-present versus absent factor as a probabilistic variable. At this point, we describe how this maps into a Bayesian model so as to see what implications can be derived for an applied theory of lineups.

Consider first a hypothetical lineup. Assume that the outcome distribution in Table 10.5 applies to the lineup. In other words, the conditions of witnessing, retention, and lineup characteristics are such that if shown a perpetrator-present lineup, we can expect a .60 probability that the witness would identify the perpetrator, a .10 probability that the witness would say “not present,” and so on. Again, from Table 10.5 assume that if the perpetrator were not in the lineup, there is a .20 probability that the witness would identify the innocent suspect, a .30 probability that the witness would say “not present,” and so on. Finally, assume that the nonlineup evidence against the suspect indicates a .50 probability that the suspect is the perpetrator and a .50 probability that the suspect is not the perpetrator. In other words, assume that it is equally likely based on the nonlineup
evidence that the witness is observing a perpetrator-present versus a perpetrator-absent lineup. How should we revise the initial .50 probability that the suspect is the perpetrator as a function of the response of the witness? Under these conditions, if the witness identifies the suspect, the revised probability becomes .75 that the suspect is the perpetrator (i.e., 60 of 80 identifications of the suspect are of the perpetrator). If the witness says “not present,” the revised probability that the suspect is the perpetrator is now only .25 (i.e., 10 of the 40 “not there” responses occur when the suspect is not the perpetrator). Similarly, if the witness picks a filler, the revised probability that the suspect is the perpetrator becomes .286 (i.e., 10 of 35 filler identifications occur when the suspect is the perpetrator). Finally, if the witness says “don’t know,” the revised probability that the suspect is the perpetrator becomes .444 (i.e., 20/45).

What happens, however, if the other (non-lineup) evidence against the suspect indicates a .70 probability that the suspect is the perpetrator rather than the .50 probability? This is where Bayesian statistics comes into play, because it articulates a simple algebra for combining probabilistic evidence. A useful form of Bayes’ Theorem links hypotheses to data with the equation

\[
p(H|D) = \frac{p(D|H) \cdot p(H)}{p(D|H) \cdot p(H) + p(D|\neg H) \cdot p(\neg H)}
\]

where \(p(H|D)\) is the probability that the hypothesis is true given the data, \(p(D|H)\) is the probability of the data given that the hypothesis is true, and \(p(D|\neg H)\) is the probability of the data given the hypothesis is not true. Now, we simply substitute lineup terms for the equation in which the data is the response of the eyewitness and the hypothesis is that the suspect is the perpetrator. To help keep this straight, we use the following terms for the witness response:

- ids = witness identified suspect
- idf = witness identified a filler
- np = witness said “not present”
- dk = witness said “don’t know”

We use the following terms for the status of the hypothesis being tested:

- \(sp\) = suspect is the perpetrator
- \(s \sim p\) = suspect is not the perpetrator

We can now describe the probability that the suspect is the perpetrator, given an identification of the suspect as

\[
p(sp|ids) = \frac{p(ids|sp) \cdot p(sp)}{p(ids|sp) \cdot p(sp) + p(ids|s \sim p) \cdot p(s \sim p)}
\]
First, we apply this equation to the situation in which the non-lineup evidence indicates a .50 probability that the suspect is the perpetrator:

\[ p(\text{sp}|\text{ids}) = \frac{(0.60)(0.50)}{(0.60)(0.50) + (0.20)(0.50)} = 0.75 \]

The .75 probability is the same as the one we calculated without the aid of Bayes’ Theorem. Now, however, what about the situation in which the other (non-lineup) evidence indicates a .70 probability that the suspect is the perpetrator (i.e., there is a 70% chance that the perpetrator is in the lineup)? In this case,

\[ p(\text{sp}|\text{ids}) = \frac{(0.60)(0.70)}{(0.60)(0.70) + (0.20)(0.30)} = 0.875 . \]

We can then apply this analysis to the other possible witness responses:

\[ p(\text{sp}|\text{np}) = \frac{(0.10)(0.70)}{(0.10)(0.70) + (0.30)(0.30)} = 0.438 \]
\[ p(\text{sp}|\text{idf}) = \frac{(0.10)(0.70)}{(0.10)(0.70) + (0.25)(0.30)} = 0.482 \]
\[ p(\text{sp}|\text{dk}) = \frac{(0.20)(0.70)}{(0.20)(0.70) + (0.25)(0.30)} = 0.651 \]

Notice how the ultimate question that confronts the trier of fact (the probability that the suspect is the perpetrator) depends critically on the conjunction of the probability that the perpetrator is in the lineup plus the probabilistic discriminating abilities of the eyewitness.

Consider, for example, what happens when we examine the data from Table 10.5 across all possible probabilities that the lineup includes the perpetrator, which is shown in Figure 10–1. The straight diagonal line is called the null line and represents the hypothetical situation in which the witness response has no diagnostic value at all. Deflections from the null line indicate some degree of diagnosticity for the witness’s response, and the greater the deflection the greater the diagnosticity of the response. Notice how the identification of the suspect produces a deflection above the null line (raising the probability that the suspect is the perpetrator), whereas the other three responses produce deflections below the line (lowering the probability that the suspect is the perpetrator).

There are a number of important observations to be made about Figure 10–1. First, Figure 10–1 illustrates how the judgment that triers of fact must make (the probability that the suspect is the perpetrator, given a particular response of the witness) is a complex, nonlinear interplay between the performance abilities of the eyewitness (assessed via
diagnosticity ratios) and the circumstance that the witness faced (perpetrator-present versus absent lineup). Second, notice that all of the curves regress toward the null line when the probability that the perpetrator is in the lineup begins to approach certainty (either 1.0, which is certainty that the lineup contains the perpetrator, or 0.0, which is certainty that the suspect is not in the lineup). This makes good sense because it means that the witness response does not much matter if one is already certain that the perpetrator is or is not in the lineup. (The reader is reminded at this point that we assume only one suspect in the lineup, and, therefore, knowing that the perpetrator is or is not in the lineup is tantamount to knowing whether the suspect is or is not the perpetrator.) Notice as well that the identifications of the suspect curve and the “not there” curve are mirror images of each other, one above the null line and the other below the null line, because both responses have diagnosticity ratios of 3.0 (see Table 10.5), albeit in opposite directions (one incriminating and the other exonerating). Readers are cautioned against assuming that these two diagnosticity ratios have to be equal. They are equal in the hypothetical data of Table 10.5, but real data show a variety of patterns that depend on witnessing and testing conditions that we do not yet fully understand (see Wells & Olson, 2002). On the other hand, there are constraints on the possible patterns that can occur in Figure 10.1, and these constraints are fully free of assumptions about witnessing and testing conditions per se. The primary constraint is that any degree of upward deflection from the null line that comes from identifications of the suspect must be accompanied by a deflection below the null line for one or more of the three nonidentification responses (i.e., “not there,” filler identification, or “don’t know”) and vice versa.

One of the useful properties of the curves in Figure 10–1 is that they specify how much “other evidence” against the suspect needs to exist in order to reach a particular level of certainty that a suspect who has been identified from the lineup is in fact the

FIGURE 10–1. Probabilities that the suspect is the perpetrator as functions of witness response and prior probability that the perpetrator is in the lineup.
Suppose, for example, one needed to be 95% certain to vote guilty. Based on Figure 10–1, this means that there would have to exist other (non-lineup) evidence indicating a .86 probability that the suspect is the perpetrator in order for the identification of the suspect by the eyewitness to push the probability past 95%. Or, if one needed to be 99% certain, then there would have to exist other (non-lineup) evidence indicating a .97 probability that the suspect is the perpetrator in order for the identification of the suspect by the eyewitness to push the probability to 99% or greater. Notice as well that if the other (non-lineup) evidence in the case indicates a 95% probability that the suspect is the perpetrator, a “not present” response from the witness lowers to .84 the probability that the suspect is the perpetrator.

An alternative way to examine the data in Figure 10–1 is to graph the deflections from the null line with the use of absolute values. The result is what has been called information-gain curves (Wells & Lindsay, 1980; Wells & Turtle, 1986; Wells & Olson, 2002), and Figure 10–2 shows these curves for the data from Table 10.5. In effect, information gain is a measure of how much the witness’s response changes the probability that the suspect is the perpetrator (regardless of direction). There are several observations about these information-gain curves that further illuminate our understanding of the complex interplay between witness abilities and the probability that the perpetrator is in the lineup. First, note that information gain curves are not symmetric. In particular, note that the identification of suspect curve is skewed to the right, whereas the “not present,” identification of filler, and “don’t know” curves are skewed to the left. Another way of describing this is to note that the identification of suspect curve peaks when the prior probability that the perpetrator is in the lineup is .37, whereas the “not present” curve peaks when the prior probability that the perpetrator is in the lineup is .63. Notice
as well that the identification of filler and the “don’t know” curves, like the “not present” curve, peak when the prior probability that the perpetrator is in the lineup is above .50. This happens because any piece of information will have more impact when it goes against the direction of the previously existing evidence than it will when it is in the same direction as the previously existing evidence. Hence, an incriminating response from the witness (an identification of the suspect) will have less informational value when the prevailing evidence indicates that the suspect is the perpetrator (i.e., \( p(sp) > .50 \)) than when the prevailing evidence indicates that the suspect is not the perpetrator (i.e., \( p(sp) < .50 \)). In contrast, an exonerating response from the witness (any of the three nonidentification responses) will have more informational value when the prevailing evidence is that the suspect is the perpetrator (i.e., \( p(sp) > .50 \)) than when the prevailing evidence is that the suspect is not the perpetrator (i.e., \( p(sp) < .50 \)). On first reflection, some readers might find this surprising. However, it simply means that new information is less informative if it agrees with what we think we already knew than if it is inconsistent with what we think we already knew.

Another observation from Figure 10–2 is that information gain diminishes and reaches 0.0 (zero) under conditions in which one already knows whether the perpetrator is in the lineup. (This is the same observation that we made about Figure 10–1, in which the curves meet the null line when \( p(sp) = 1.0 \) or \( 0.0 \).) This is just another way of saying that the result of the lineup (i.e., the witness response) is irrelevant if we already know with certainty from other evidence that the suspect is the perpetrator (for example, because of a DNA match).

There is one more important observation that needs to be made from Figure 10–2. In particular, we draw the reader’s attention to the fact that the incriminating curve (i.e., identification of suspect) and two of the exonerating curves (“not there” curve and filler identification curve) cross at some point. For expository purposes, consider the identification of suspect curve and the filler identification curve. On average, an identification of suspect is more informative (of guilt) than is a filler identification (of innocence), as evidenced by the higher curve for identifications of suspect than for filler identifications. (This can be noted from Table 10.5 as well because the diagnosticity ratio for identifications of suspect is 3.0, whereas the diagnosticity ratio for filler identifications is only 2.5.) However, note that when the probability that the perpetrator is in the lineup exceeds .60, then a filler identification is more informative of innocence than an identification of the suspect is of guilt.

Although theories can be couched at many different levels of analysis and abstraction, we cannot imagine an adequate theory of lineups that does not provide a representation of the complex relation between the abilities of eyewitnesses (assessed via diagnosticity ratios) and the probability that the lineup includes the actual perpetrator.

**PATTERNS OF RESPONDING**

Clearly, there are many variables that affect eyewitness identification performance. It is not the purpose of this chapter to review all of these variables. What is important to
note, however, is that eyewitness identification data tend to be patterned, and we note here the general nature of that pattern as it relates to the notion of memory strength. By memory strength, we mean the extent to which the witness has a reasonably accurate and accessible memory of the perpetrator. Clearly, memory strength is going to be determined by a host of factors at the time of encoding (e.g., distraction, arousal, distance, lighting) and during the retention interval (e.g., passage of time, post-event information, rehearsal). For our purposes, we refer simply to strong versus weak memories.

Eyewitness identification data tends to follow a predictable pattern in the literature. First, identifications of the perpetrator from the perpetrator-present lineup tend to be more probable than are identifications of the innocent suspect from the perpetrator-absent lineup. Second, “not present” responses are more probable from perpetrator-absent lineups than from perpetrator-present lineups. Third, identifications of fillers are somewhat more probable from perpetrator-absent lineups than from perpetrator-present lineups. Finally, “don’t know” responses are slightly more probable from a perpetrator-absent lineup than from a perpetrator-present lineup. The first two patterns, concerning identifications of the suspect and ‘not present” responses, are precisely what would be expected from above-chance performance of eyewitnesses. Indeed, Clark (2003) notes that the greater the difference obtained from perpetrator-present and perpetrator-absent lineups in identification of suspect rates, the greater the difference in “not present” responses. This pattern can easily be understood from a consideration of the role of memory strength. As memory strength increases, the rate of identification of the perpetrator from a perpetrator-present lineup should go up and the rate of identification of the innocent suspect from a perpetrator-absent lineup should go down. Conversely, as memory strength increases, the rate of correct rejections should increase and the rate of false rejections should decrease. In terms of diagnosticity, increases in the diagnosticity of identifications of the suspect are associated with increases in the diagnosticity of “not present” responses.

The pattern of filler identifications, on the other hand, is more complex. Clearly, the overall rate of filler identifications should decrease with increasing memory strength. But how does the differential rate of filler identifications in the present versus absent lineup (diagnosticity) vary as a function of changes in memory strength? We know of no data that have specifically addressed this issue, but based on Wells and Olson’s (2002) notion of why filler identifications are diagnostic, we would expect the diagnosticity of filler identifications to increase with increasing memory strength. Although filler identifications themselves should become less frequent with increased memory strength, the ratio of filler identifications in absent versus present lineups should increase with increasing memory strength.

THE PLEADING EFFECT

The previous sections make clear one reason why it is difficult to simply go from an experiment to the real world in estimating the probability that a suspect who is identified is in fact the culprit; such estimates depend critically on the base rate proportion of line-
ups that contain the actual culprit. But there is another factor that also has to be considered when one is trying to generalize to trials. A trial occurs relatively late in the process. By the time of trial, various prior events are likely to have differentially affected the proportions of guilty versus innocent defendants. One especially important factor is the occurrence of guilty pleas and the striking of plea bargains (see also Penrod, 2003). Most estimates place the guilty plea (plus plea bargain) figure at 80–90% (Cole, 1986). This pretrial pleading has the effect of changing the proportions of innocent and guilty persons who choose to go to trial, an effect we call the **pleading effect**.

How does the pleading effect work, and how much could it change the probabilities that an eyewitness in a trial might have made a mistaken identification? Let us assume for purposes of exploration that 85% of guilty individuals charged with a felony plead guilty, thus choosing to not go to trial. Although we concede that false confessions occur and that innocent people can sometimes plead guilty (see Kassin, 1997), let us assume for the moment that all who plead guilty are in fact guilty. Consider now that the population of those who choose to go to trial and persist with a not guilty plea is a mix of guilty persons and innocent persons. However, when these go-to-trial proportions are compared back with the proportions of identified suspects who were mistakenly identified, the proportions are very different. Suppose, for example, out of 10,000 suspects identified from lineups, 95% (9,500) are in fact guilty and only 5% (500) are innocent. If 85% of guilty suspects plead guilty, and thereby do not proceed to trial, then 1,425 of the guilty suspects would go to trial. If none of the innocent suspects plead guilty, then all 500 of the innocent suspects would proceed to trial. This means that, among the 1,925 suspects at trial, 500 (26.0%) would be innocent. Hence, what was only a 5% rate of mistaken identification at the time of the lineup becomes a 26% rate of innocence at the trial level.

We are not suggesting that the proportion of innocent defendants in eyewitness identification cases who plead not guilty and go to trial is 26%. Obviously, this figure depends, among other things, on the original proportion of identifications of suspects that are mistaken. Although we cannot know precisely what these rates are, the pleading effect pattern can be described as a curvilinear relation between the mistaken identification rate at the lineup and the innocence rate among suspects who proceed to trial. Figure 10–3 graphs this curve for lineup identification error rates from 0.0% to a rate of 20.0%, using an 85% pleading effect. Notice how the 85% pleading effect manages to make even quite low rates of mistaken identification at the lineup result in surprisingly high rates at trial. For example, 2.0% mistaken identifications at the lineup result in 12.0% of the suspects who proceed to trial being innocent.

There are two complementary ways of looking at the pleading effect. First, the pleading effect means that even slight increases in the error rate for suspect identifications at the lineup (e.g., an increase from a 1.0% rate to a 3.0% rate) produce larger increases in the proportion of those who go to trial being innocent (an increase from a 6.3% rate to a 17.1% rate). At the same time, this means that improvements to lineups that yield seemingly small reductions in the rate of mistaken identification (e.g., a decrease from a 4.0% rate to a 2.0% rate) will yield considerable reductions in the proportion of those who go to trial being innocent (a decrease from a 21.7% rate to a 12.0% rate).
PSYCHOLOGICAL PROCESSES: A BROAD VIEW

Undoubtedly, psychological researchers will find more to quibble with in this section on psychological processes than in the previous sections (on the function of lineups, the structure of lineups, and outcome distributions). Indeed, it is here that we offer considerably more conjecture and it is here where we expect future experiments to highlight the need for revising our ideas. What we describe here are broad psychological processes that we presume operate at the time of the lineup. Critics might say that we are using the term broad as a euphemism for loosely defined or ambiguous. Maybe such critics would be correct in this assertion, because what we really mean is that we believe that something like these processes seems to be operating, and, although we might not be able to distinguish between subtle variations on these presumed process or their interactions, they seem to have some explanatory power.

Note as well that we are not describing a theory of memory acquisition or memory storage. For example, our lineup theory does not attempt to address such things as why cross-race identifications are less reliable than within-race identifications, the role of stress at the time of witnessing, and so on. Instead, our discussion is focused primarily on retrieval processes that operate at the time of the lineup.

FIGURE 10–3. Percent of those proceeding to trial who are innocent as a function of mistaken identification rates at lineup using an 85% pleading effect.
There has been relatively little empirical work investigating the psychological processes that operate at the time of lineup presentation; what has been done has generally posited a dual processing framework (e.g., Dunning & Stern, 1994). Dual-processing theories in social and cognitive psychology have proliferated, especially over the last decade (Chaiken & Trope, 1999; Sloman, 1996). Dual-processing theories posit that there are two systems by which people process information. Specific characterizations of these systems vary from one theorist to another, but they tend to argue that one system is largely effortless, rapid, and holistic, whereas the other system is relatively effortful, slow, and analytic. Shiffrin and Schneider’s (1977) distinction between automatic and controlled processing is one example. Others include Petty and Cacioppo’s (1981) central versus peripheral processing model, Chaiken’s (1980) heuristic versus systematic processing distinction, Epstein’s (1994) experiential versus rational processing distinction, and Fazio and Towles-Schwen’s (1999) spontaneous versus deliberative processing distinction. Dual-processing theories vary in certain respects, such as the extent to which one type of processing is more conscious than the other, the extent to which the two processing systems are exclusive versus non-exclusive, and so on, but they all share the same basic distinction between the two processes.

Similar to these theories, Dunning and Stern (1994) originally distinguished between two processes in which witnesses may engage when presented with a lineup: automatic recognition and process of elimination. According to this conceptualization, automatic recognition judgments tend to be quick and effortless and are reached with little conscious strategy. Process of elimination judgments (otherwise known as deliberative judgments), in contrast, tend to be slow and effortful and are reached through deliberate and conscious strategies.

The automatic versus deliberative distinction can be partially appreciated by performing the two tasks shown in Figure 10–4. The task in the top panel is to decide whether the face is that of former President Bill Clinton. The task in the bottom panel is to decide which person is the suspect in the lineup. Consider the differences in how these two tasks are performed. In the case of deciding whether the picture was of Bill Clinton, the decision process was likely to be very rapid, effortless, and holistic. If asked how it was done, people would largely be at a loss for words, perhaps saying something like “I don’t know. I just recognized him,” and there would be little awareness of the complex cognitive processes that were involved. Even if a person were performing another task at the same time (e.g., talking on a phone), the Clinton task could be performed because it consumes no significant cognitive resources to perform that task. This is what is meant by an automatic process. Consider now the “which person is the suspect?” task. Compared with the Clinton task, this task took more time, involved some effort, and was performed at a particularistic level. People tend to be consciously aware of the process by which they made the judgment, and, if asked how it was done, people could probably verbally articulate some type of reasoning they used. This is what is meant by deliberative.
Dunning and Stern (1994) argued that because a lineup is basically a recognition task (an eyewitness must state whether he/she recognizes the criminal from among the members of the lineup), accurate lineup identifications should tend to be the result of automatic recognition judgments. In contrast, inaccurate identifications, which almost by definition are identifications of people the eyewitness has never seen, cannot be the result of a recognition judgment and should therefore tend to be the result of a deliberative process. Thus, by determining the strategy that eyewitnesses used to arrive at their identification decision, one should show some ability to differentiate accurate from inaccurate witnesses. Consistent with predictions, Dunning and Stern, as well as others since them (e.g., Kneller, Memon, & Stevenage, 2001; Lindsay & Bellinger, 1999), have found that when witnesses to a mock crime are asked after having made an identification about their decision processes, accurate witnesses are more likely to endorse having engaged in automatic processes (e.g., saying that the culprit’s photo “popped out” from the array), whereas inaccurate witnesses are more likely to endorse having engaged in deliberative processes (e.g., saying that they compared photos with other photos or used methods of elimination).

More indirect evidence supports the automatic/deliberative judgment distinction. Identification accuracy is harmed by forcing of eyewitnesses to rely more on deliberate, verbal processes, whether by having them verbally analyze features of a criminal’s face at the time of the crime (Wells & Turtle, 1988), by having them verbally give a description of a criminal prior to the lineup presentation (Schooler & Engstler-Schooler, 1990), or...
by having them verbally describe reasons why each person in the lineup might or might not be a good match to the criminal (Perretta & Dunning, 2001). Consistent with this, Schooler and Engstler-Schooler found that limiting the amount of time a witness had to make an identification to 5 seconds eliminated the negative effects of verbalization, presumably because those witnesses were not given enough time for the verbalization to produce its deleterious effects on accuracy.

Multiple studies also have shown that the amount of time it takes for a witness to make an identification is negatively correlated with accuracy; the quicker the decision, the more likely it is to be correct (Dunning & Stern, 1994; Smith, Lindsay, & Pryke, 2000; Sporer, 1992, 1993, 1994; Weber & Brewer, 2003). This suggests that accurate witnesses are the ones who tend to experience quick, automatic recognition experiences, as opposed to engaging in slower, deliberate processes. Finally, consistent with the idea that automatically made judgments should be less affected by contextual factors, such as array size, than are deliberately made judgments, Perretta and Dunning (2001) found that increasing the lineup size increased response latency more for inaccurate than for accurate witnesses. Thus, inaccurate witnesses seemed to be relying more on deliberative judgments than accurate witnesses.

The distinction between automatic and deliberative processes is similar to the distinction between absolute and relative judgments (Wells, 1984; Lindsay & Wells, 1985). An absolute judgment is one in which a witness compares individual members of the lineup with his or her memory of the criminal. A relative judgment, in contrast, is one in which a witness compares members of the lineup in order to determine who looks most like his or her memory of the criminal. It has been found that witnesses who report engaging in an absolute judgment strategy are more accurate than witnesses who report engaging in a relative judgment strategy (Lindsay & Bellinger, 1999; Smith, Lindsay, & Pryke, 2000), and that forcing witnesses to adopt an absolute judgment strategy decreases false identifications (Lindsay & Wells, 1985). Although it has never been empirically tested whether absolute and relative judgments are automatic or deliberate processes, it has often been assumed that absolute judgments are automatic processes and relative judgments are deliberate processes; in fact, the terms have often been used interchangeably in the literature. This is, to a certain extent, understandable. Absolute judgments, like automatic judgments, are memory-dependent; both require a relatively strong memory of the criminal and should be only minimally influenced (if at all) by superficial features of the lineup. Witnesses who report having used absolute judgments, like witnesses who report having had experiences of automatic processes, tend to be more accurate in their identification decisions. In contrast, the relative judgment strategy of comparing lineup members to other lineup members is an effortful, time-consuming, and deliberate process that relies less on memory than on differences between lineup members. The tendency for witnesses to engage in relative judgments can be reduced through specific instructions (Malpass & Devine, 1981), suggesting that relative judgments are under some form of deliberate control. In fact, the similarities between absolute/relative judgments and automatic/deliberative processes can be so striking that some of the self-report questions used to determine the extent to which automatic and deliberative processes were operating during the identification procedure are the same ones used to determine
the extent to which the witness used an absolute versus relative judgment strategy (e.g.,
the extent to which witnesses endorse the statement *I compared the pictures to each other to
make my decision* is a measure of both deliberative processes and relative judgments).

It is, however, empirically tenuous to conclude that absolute judgments are necessarily automatic and relative judgments are necessarily deliberative. Absolute judgments, despite their reliance on memory, are nonetheless a process of comparison (of each lineup member to one’s memory). The extent of this comparison is probably under the witness’s control (how superficially the comparisons are made, the speed with which each comparison is made, whether the comparison is based on featural or holistic qualities, etc.). To the extent that absolute judgments are under the witness’s control, they cannot be completely automatic. Similarly, relative judgments are not easily classifiable as deliberate judgments. One can imagine relative judgments being broken down into three parts: comparing lineup members to each other, using that comparison to determine who from the lineup looks most like the criminal, and deciding whether the person in the lineup who looks most like the criminal actually is the criminal. The first part, for reasons stated above, is probably largely deliberative. However, the next two parts are not as clearly deliberative processes. It may be, for example, that one’s determination of how good a match X is to Y is an automatic process unavoidably resulting from a process of comparison. Or perhaps deciding whether the best match in the lineup actually is the criminal can be an automatic process and an unavoidable consequence of determining the best match to one’s memory. Thus, although we believe that on average absolute judgments tend to be more automatic and relative judgments tend to be more deliberative, we urge researchers to be cautious in their assumptions as to the cognitive processes underpinning absolute and relative judgments.

Although there is evidence supporting the distinction between automatic and deliberative judgments, there also appears to be evidence that questions and challenges this simple dichotomy. Specifically, we advance the idea that automatic and deliberative judgments, instead of being dichotomous, actually lie on opposite ends of a continuum of judgment processes, and that instead of cognitive processes actually causing identification accuracy per se, there is a third variable—the quality of one’s memory of the criminal—that drives both accuracy and judgment process and accounts for their covariation. Note that we are not rejecting the idea that automatic and deliberative processes are operating during a lineup identification; in fact, we embrace the distinction and think it has much to offer. Instead, we are suggesting potential refinements to the automatic/deliberative conceptualization that we think have solid theoretical grounding and that help explain some of the otherwise anomalous data. We will now examine the basis of our claims.

**DOES DECISION PROCESS CAUSE DIFFERENCES IN ACCURACY?**

The basic mechanism that is presumed to account for the correlation between self-reported cognitive process (automatic or deliberative) and accuracy is as follows: A witness is
shown a lineup that either contains the actual criminal or does not contain the actual criminal. If it does contain the actual criminal, the witness should tend to undergo an immediate sense of recognition of the criminal; the criminal “pops out” of the lineup, and the witness will make a quick and accurate identification. If, however, the criminal is not in the lineup, and thus the witness does not experience that immediate sense of recognition, deliberative processes kick in. The witness will compare pictures and engage in other sorts of deliberate and cognitively effortful behaviors that will often eventually lead the witness to make a slow and necessarily inaccurate identification. Thus, quick and automatic decisions tend to be accurate, whereas slow and deliberate decisions tend to be inaccurate.

This explanation makes an easily testable prediction, namely, that interfering with a witness’s ability to engage in deliberative processes and forcing witnesses to rely on automatic judgments should decrease the likelihood of that witness making a false identification. There are two ways in which this prediction has been tested. One way is by giving witnesses a cognitive task to perform while making a lineup identification, which should interfere with deliberative processes but not with automatic processes (Kahneman & Treisman, 1984). Being unable to engage in deliberative cognitive processes should force witnesses to rely on automatic judgments, and this manipulation should therefore increase witnesses’ accuracy. Following this logic, Perretta and Dunning (2001) report a study in which some of their participant-witnesses were given a 9-digit number to memorize while attempting an identification from a lineup, whereas other participants were not given a number to memorize while attempting an identification. The authors predicted that the witnesses who were given the cognitive task of memorizing a number would not be able to engage in harmful, deliberative processes while viewing the lineup and should hence be more accurate than witnesses who were not given this task. Contrary to this prediction, those who were made cognitively busy were not significantly more accurate than controls.

The other way in which experimenters have interfered with deliberative processes is by limiting the amount of time witnesses have to make a lineup identification, which should necessarily limit the amount of deliberative processing in which the witnesses can engage. Thus, imposing time constraints should force eyewitnesses to rely more on automatic judgments than on deliberative judgments, which should increase their accuracy. In fact, before the automatic/deliberative judgment distinction even surfaced in the eyewitness area, Schooler and Engstler-Schooler (1990) had data that spoke to this prediction. Although they were not testing this prediction directly, one of their studies involved limiting the amount of time people had to make facial recognition judgments to 5 seconds. When witnesses gave a description of the target before viewing the lineup, which was found to impair accuracy, this time constraint eliminated the deleterious effects of giving the description. However, when witnesses did not give a description of the target before viewing the lineup, the time constraint had no significant effect on accuracy. Following similar protocols, other researchers have also failed to find significant effects of limiting decision time on accuracy rates (Brewer, Gordon, & Bond, 2000; Charman, 2004; Perretta & Dunning, 2001).
Procedures that inhibit deliberative processing do not seem to cause an increase in eyewitness accuracy. On the other hand, as previously mentioned, there is some evidence that procedures that facilitate deliberative processing cause decreases in eyewitness accuracy (e.g., Perretta & Dunning, 2001; Schooler & Engstler-Schooler, 1990; Wells & Turtle, 1988).

How can facilitating deliberative processing cause decreases in eyewitness identification accuracy, whereas inhibiting deliberative processing does not cause increases in eyewitness identification accuracy? One possibility is that past studies showing a negative effect of increased deliberation on accuracy have confounded increased deliberation with memory interference. Specifically, researchers who have attempted to increase deliberative processes have done so by giving their participants an additional task to perform, such as verbalizing some aspect of the criminal or of the lineup procedure. It may be, therefore, that it was interference of the witnesses’ memories of the criminal caused by the additional task that caused the detrimental effects. For example, perhaps a witness who gives a verbal description of a criminal prior to a lineup becomes biased to retrieve only description-consistent (i.e., verbalizable) information about the criminal from memory at the time of the lineup task and to ignore other, more diagnostic information about the criminal, leading the witness to make an inaccurate identification. Thus, it might not have been the increase in deliberative processing per se that caused the detriment in accuracy, but rather the interference produced by the additional task of verbalizing the criminal’s face. Therefore, deliberative processing, whether increased or decreased, may actually have no causal effects whatsoever on eyewitness accuracy; the apparent harm in accuracy caused by increased deliberative processing may actually reflect the harmful effects of interference.

QUALITY OF MEMORY AS A THIRD VARIABLE

If the type of decision process in which witnesses engage does not cause accuracy, how then can we explain the correlation between self-reported decision process and accuracy? There are two possibilities. The first is that differences in accuracy actually cause differences in self-reported decision processes. Consistent with this idea, studies have found that leading eyewitnesses to believe that they had made an accurate identification causes them to report that their process was more automatic (e.g., increase their reports that the identified person’s photo “popped out” from the array; e.g., Bradfield, Wells, & Olson, 2002; Wells & Bradfield, 1998). In other words, it seems that eyewitnesses’ beliefs about their accuracy influenced self-reported processes. However, this explanation seems to be lacking. It seems unlikely that identification accuracy per se is a sufficient cause of the differences in retrospective self-reports. What is the mechanism through which accuracy affects self-reports? What is different about accurate and inaccurate witnesses that would lead to the differential self-reports of cognitive decision process?

The answer to this question, and the second possible explanation of the correlation between decision process and accuracy, is that there may exist a third variable that influences both decision process and accuracy. This idea has been offered by Brewer et al.
(2000), who suggested that the quality of the memory that a witness has of the criminal may be this third variable. To examine this idea more closely, imagine that a witness who has a poor memory of the criminal is asked to attempt an identification from a lineup. It seems reasonable to assume that, because she has a poor memory of the criminal, she is less likely to make a correct decision. But how does the quality of the memory lead to differences in decision process? Let us revisit the absolute/relative judgment strategy distinction. As previously mentioned, absolute judgment strategies are more dependent on memory than relative judgment strategies because absolute judgments require comparisons between individual lineup members and the witness's memory, whereas relative judgments simply require comparisons among the various lineup members. Thus, witnesses who tend to use more absolute judgment strategies should be those who have a relatively strong memory for the criminal. Let us return to our witness who has a poor memory of the criminal. Because of her lack of a strong memory of the criminal, she may therefore be tempted to forgo an absolute judgment strategy and instead engage in a relative judgment strategy, something that witnesses readily do (Wells, 1993). Thus, because of her poor memory, she will tend to be inaccurate and will tend to engage in a relative judgment strategy. Contrast this with a witness who has a strong memory of the criminal. Because of the strength of the memory, this witness is likely to be accurate in her identification. Also because of the strength of the memory, this witness can now engage in an absolute judgment strategy to make her decision. Thus, witnesses with strong memories are more likely to be accurate witnesses and should also tend to engage in different cognitive decision processes than would inaccurate witnesses.

Can this third variable explanation account for the data? We contend that it does. First, this explanation accounts for the negative correlations found between decision time and accuracy. Imagine that decision time is a function of the number of comparisons that are made from a lineup. When using an absolute judgment strategy, witnesses compare each lineup member with their memory of the criminal; thus, the number of comparisons to make is simply equal to the number of members in the lineup. When using a relative judgment strategy, however, witnesses compare people in the lineup with other people in the lineup; thus, the number of comparisons to make can be as high as

$$\sum_{i=1}^{n} (n - i)$$

where $n$ is equal to the number of members in the lineup. For example, with a six-person lineup, an absolute judgment strategy would take up to six comparisons, whereas a relative judgment strategy would take up to $5 + 4 + 3 + 2 + 1 = 15$ comparisons. Thus, relative judgments should tend to be slower than absolute judgments, and because relative judgments tend to indicate a poorer memory of the criminal, they should also tend to be more inaccurate than absolute judgments. Therefore, this third variable approach would predict that decision time should be negatively correlated with accuracy, without invoking the idea that automatic/deliberative processes causally determine decision time.

Second, the quality-of-memory-as-third-variable explanation accounts for both the interference explanation of the detrimental effects of having witnesses verbalize features...
of the criminal as well as the null effects of encouraging witnesses to rely on automatic processes. Once a witness has encoded a memory of the criminal, it is unlikely that the quality of that memory can improve, but it is very likely that the quality of the memory can be degraded by various manipulations. Experimental manipulations that lead witnesses to verbalize their memories of a criminal’s face may ultimately degrade the quality of the memories that are accessed during lineups by leading witnesses to selectively retrieve features of the criminal’s face that were expressed verbally and to ignore more diagnostic, nonverbally expressed features. This degraded memory would then tend to lead witnesses to engage in deliberative judgment processes and to inaccurate decisions. Thus, the observed detriment in accuracy resulting from increased deliberation about the lineup may not be a result of a change in decision process per se, but may instead originate through the quality-of-memory variable. However, because the quality of a memory of a face cannot improve over time (without additional exposure to that face), attempts to increase accuracy beyond a baseline should be doomed to fail. This would explain why limiting deliberative processing does not increase accuracy. The effects of encouraging and discouraging deliberate processing can therefore be explained without resorting to the idea that decision process differences cause changes in accuracy; rather the findings can be explained as the result of a third variable causing both decision process differences and changes in accuracy.

Finally, this explanation accounts quite easily for the differential self-reports of decision strategy between accurate and inaccurate witnesses. Because accurate witnesses tend to be those who have a good memory of the criminal, and those who have a good memory of the criminal tend to be those who use quicker and more automatic judgments, accurate witnesses should be more likely to endorse having used absolute judgment strategies. Similarly, because inaccurate witnesses tend to be those who have a poor memory of the criminal, and those who have a poor memory of the criminal tend to be those who use more relative judgment strategies, inaccurate witnesses should be more likely to endorse having used deliberative judgment strategies. Although we argued earlier that absolute and relative judgments are not necessarily automatic and deliberative processes, respectively, we nonetheless believe that an absolute judgment strategy tends to be more automatic than a relative judgment strategy. Thus, witnesses who use an absolute judgment strategy should be more likely to be accurate and to report having engaged in an automatic process than witnesses who use a relative judgment strategy.

A CONTINUUM OF AUTOMATICITY

The idea that the quality of memory drives both accuracy and decision processes raises an interesting conundrum. Because quality of one’s memory is a continuous variable, how is it that it leads to a dichotomous automatic/deliberative decision process? Does the variability in memory quality get ignored in the automatic/deliberative distinction? We contend that the answer is no. We believe that labeling many processes as simply deliberative or automatic obfuscates finer gradations of automaticity. Instead of treating the
automatic/deliberative distinction as a dichotomy, it may be more correct to think of
cognitive processes as lying somewhere along an automaticity continuum, from low
automaticity to high automaticity. According to this view, deliberative processes will
dominate to the extent that automatic processes are toward the low end of the auto-
maticity continuum. Thus, a lineup task may not be simply a matter of automatic versus
deliberate processes, but may rather be a matter of relatively automatic versus relatively
deliberate processes. Let us now examine the evidence for this assertion.

There are at least five reasons why a simple automatic/deliberative process di-
chotomy may be inaccurate. The first two are theoretical reasons why purely automatic
recognition responses from eyewitnesses should be difficult to develop, and the subse-
quent three are reasons based on empirical evidence that speak against a strict dichotomy.
First, from a cognitive psychology standpoint, the development of automaticity generally
takes repeated and/or prolonged exposure to a stimulus (Shiffrin & Schneider, 1977). Al-
though a prolonged exposure to a criminal is certainly possible in many cases (e.g., kid-
nappings, hostage situations), most real-world eyewitness probably do not receive the ex-
posure they need to develop a purely automatic recognition response.

Second, automatic processes are very stimulus-dependent (Shiffrin & Schneider,
1977). That is, the stimulus that developed the automatic response in the first place is
required to instantiate it later. Automatic responses do not generalize very well to novel
or changed stimuli. Thus, an automatic recognition judgment would be dependent on
the match between the physical appearance of the criminal at the time of the crime and
the physical appearance of the criminal at the time of the lineup. However, there is little
reason to assume that this match would be very strong. Pictures used in lineups often are
not taken close to the time of the crime, and thus the criminal is likely to have changed
appearance. Without a strong similarity between the criminal at the time of the crime
and the criminal at the time of the lineup, purely automatic responses are unlikely.

Third, what are considered automatic judgments in the eyewitness literature often
take much longer than typical cognitive automatic judgments. For example, Dunning and
Perretta (2002) found across four studies that a decision time of 10–12 seconds maximally
differentiated accurate from inaccurate witnesses and, by implication, those who used au-
tomatic and deliberative processes, respectively. Pure automatic recognition judgments,
however, are typically much quicker. Imagine identifying a photo of your own face from
a lineup; that process should be virtually instantaneous. Although it could be argued
that Dunning and Perretta’s participants who did experience automatic recognition all
responded within, say, 2 seconds, one would then be forced to conclude that participants
who took 2 to 12 seconds to respond used deliberative processing and were much more
accurate than those who used deliberative processes and took longer than 12 seconds to
respond. This would not only break down the dichotomy, but would also mean that de-
liberative processes are often highly accurate, a position antithetical to the automatic/
deliberative distinction. It also could be argued that Dunning and Perretta’s accurate
participants did experience an automatic judgment, but looked at the other pictures to
verify their experience, thus increasing their decision time. This argument is supported
by the finding that accurate witnesses were more likely than inaccurate witnesses to
endorse the statement *They [the pictures] helped me to confirm, reinforce my decision after I made it.* However, this only indicates that accurate witnesses actually used a combination of automatic and deliberative judgments in making their decision, a position more compatible with a continuum idea of automaticity than a strictly dichotomous view. A purely automatic response would require no comparative judgments between pictures.

Fourth, the boundary that maximally differentiates accurate from inaccurate witnesses is not stable across experimental variations and has been found to vary from as little as 5 seconds to as much as 29 seconds (Weber, Brewer, Wells, Semmler, & Keast, 2004). If accurate responses to lineups truly are a function of pure automatic responding, they should be relatively insensitive to experimental variations. The variability in the boundary that best differentiated accurate from inaccurate responses suggests that accurate decisions are not completely automatic, because automatic recognition processes should be invariant to contextual changes (Treisman & Gelade, 1980).

Fifth, an experiment by Dunning and Perretta (2001) showed that increasing lineup size increased decision time more for inaccurate than for accurate witnesses, suggesting that inaccurate witnesses, more than accurate witnesses, were engaging in deliberative processes such as comparing pictures. Although we agree with this interpretation, we do not think that this indicates that accurate witnesses were relying on purely automatic judgments. A purely automatic judgment should not be affected at all by an increase in array size (Schneider & Shiffrin, 1977). In fact, Dunning and Perretta found that increasing lineup size increased the time it took accurate witnesses to make a decision, suggesting that accurate identifications are not a function solely of automatic judgments. It can be argued that because accuracy is only probabilistically associated with automaticity, perhaps some accurate witnesses actually made their decisions through a deliberative process of comparing pictures, which would create an effect of lineup size on decision time for accurate witnesses. We agree that this is possible, but think it is at least as possible, if not more so, that even accurate eyewitnesses used some deliberative processing when making their decisions.

Note that a continuum of automaticity is consistent with theory and data. The degree to which a response to a lineup is automatic would be dependent on the duration of exposure a witness had to the criminal, as well as the degree to which the physical appearance of the criminal at the time of the crime matches the physical appearance of the criminal at the time of the lineup. Because these variables would not be constant across multiple experiments, different experiments should lead to differing degrees of automaticity and hence differing maximally differentiating time boundaries. This time boundary would be longer than that for purely automatic responses, to the extent that witnesses were not using purely automatic processes. Similarly, witnesses’ decision times should be influenced by lineup size to the extent that they were not using purely automatic processes. Thus, Dunning and Stern’s (1994) finding that accurate witnesses endorsed having used more automatic judgment strategies than inaccurate witnesses does not necessarily indicate that accurate judgments are automatic per se, only that they are relatively more automatic than inaccurate judgments.
We have so far laid out two main arguments for how we think the automatic/deliberative distinction can be improved: We have argued that a third variable—quality of memory for the criminal—may be driving both accuracy and decision process, and we have argued that the automatic processes may be better thought of as a continuous variable. Is it possible to integrate these new ideas into a cohesive theory? This is the question to which we now turn.

A COMPETITION/CORROBORATION CONCEPTUALIZATION OF LINEUP IDENTIFICATION

We propose a theoretical conceptualization of lineup identification that is borrowed in part from Logan's (1988) instance theory of automatization, a popular and influential theory within the field of cognitive psychology. Because Logan's theory was developed largely within the context of other types of tasks (e.g., performing math problems), however, we needed to make certain modifications to account for the fact that lineups involve visual recognition memory rather than recall memory. Although instance theory has more recently been updated to include more general attention and memory processes (Logan, 2002), we describe the original version of the theory, which deals specifically with automatization and is thus most relevant.

Instance theory was proposed as an alternative account of the generally held ideas of the nature of automatic processes. Whereas earlier explanations of automaticity relied on the premise that the development of automaticity was equivalent to the gradual withdrawal of attention, instance theory regards the development of automaticity as a memory phenomenon that reflects a transition in reliance on cognitive processes. According to the theory, one starts out by responding to a novel stimulus with slow and tedious algorithmic responses, but with practice one accumulates discrete memories, or instances, of interacting with and responding to the stimulus. Eventually, as these instances accrue, one becomes able to respond to the stimulus on the basis of memory for those instances and abandons reliance on algorithms. Importantly, however, algorithmic responding and memory-based responding do not operate exclusively of one another; rather they always operate in parallel. Logan (1988) referred to their parallel operations as a race between the two processes, whereby the one that finishes first controls the response. Thus, as instances accumulate over time, memory-based responding is more likely to win the race over algorithmic-based responding, simply because there are more memories from which to draw. For our purposes, we consider algorithmic and memory-based processes to be roughly analogous to deliberative and automatic processes, respectively.

A number of points are important to note about this framework. First, the degree to which a lineup response is automatic or deliberative is dependent on the quality of the memory (the number of instances) that the witness has of the criminal such that the better the memory, the more automatic the response. Thus, the quality of memory not only determines accuracy, but also the cognitive process that is undertaken. Second, because
automaticity is thought of as the relative quickness of memory-based judgments over algorithmic-based judgments, which in turn depends on the number of instances one has in memory, automaticity surfaces as a continuous variable. In fact, in Logan's (1988) view, no process can become completely automatic, because the addition of any single instance would increase the degree of automaticity (albeit with increasingly diminishing returns). Therefore, we can refer to cognitive processes in terms of degree of automaticity as opposed to simply automatic versus not automatic. Additionally, we can see how the degree of automaticity experienced is moderated by various factors. Anything that affects a witness's memory for the criminal (e.g., exposure time, match between criminal at the time of the crime and criminal at the time of the lineup) should affect the degree to which that witness reports experiencing a process of automaticity. Third, because strong memories are associated with both accuracy and the use of quick, memory-based processes, quick identification decisions will tend to be accurate. Fourth, memory-based and algorithmic processes are assumed to work in parallel. This is important in that eyewitness researchers have usually assumed that when automatic processes are operating, deliberative processes are not operating, and vice versa. It may be that witnesses can have an automatic process operating simultaneously with a deliberative process.

Differences between Our Competition/Corroboration Conceptualization and Instance Theory. Our conceptualization differs from Logan's (1988) Instance Theory in a few key respects. First, instead of construing of automaticity as something that accrues as a function of discrete instances of interacting with the stimulus, our conceptualization assumes that automaticity can accrue through any number of processes that increase the quality of the eyewitness’s memory for the stimulus. Although repeated instances of exposure should increase automaticity, automaticity should also increase with exposure duration within any given instance, the degree of attention paid to the stimulus, the quality of the view that the witness had during any given instance, and so on. Second, because the original exposure(s) and the lineup test are complex visual stimuli (rather than, for example, math problems), the degree of automatic response to the perpetrator in a lineup should depend not only on the quality of the original memory, but also on the extent to which the perpetrator’s physical appearance in the lineup resembles the perpetrator’s appearance at the time of the crime. A math problem is essentially the same whether it is presented as “9 × 9 = ___” or as “nine times nine equals ___?” A visual identification test, however, is heavily dependent on the extent to which the visual stimulus is a replica of the original exposure. In our conceptualization, this means that an automatic response would normally be nonexistent when the actual perpetrator is not in the lineup. We acknowledge, however, that a weak automatic response could occur if there were a person in the lineup who was highly similar to the perpetrator. Third, whereas automatic and deliberative processes are construed in Logan's theory as being involved in a race of time, our conceptualization presumes that automatic processes will always win a time race. Instead of competing in a time race, our conceptualization assumes that when an automatic response is experienced, deliberative processes are usually engaged in as well before any decision is made by the eyewitness. In other words, even when an automatic process is experienced, we propose that eyewitnesses nevertheless engage in
deliberative processes in an attempt to either corroborate or call into question the validity of the automatic response. This use of a deliberative process is almost certain to occur in the absence of an automatic process and is very likely when the automatic experience is weak, but even a strong automatic response is likely to be followed by deliberation. The amount of deliberation that is needed is diminished by the strength of the automatic response (which accounts for faster decision times for accurate than for inaccurate witnesses). Nevertheless, some amount of deliberative processing is probably always involved. The use of deliberative processes when there is an automatic response stems at least in part from the importance of the decision (emanating from such considerations as the importance of not making a mistake and consequences for the witness and the accused). Furthermore, the absence of time constraints on the decisions of eyewitnesses helps guarantee that deliberative processes will almost always be engaged. Finally, we propose that the deliberative processes that follow an automatic response will generally be biased toward confirmation of the automatic response. Nevertheless, there will be circumstances in which deliberative processes conflict with automatic processes and thereby compete for the decision. For example, an automatic reaction to lineup member number 2 might be followed by the witness noting that lineup member number 4 is the only photo with a different color background, possibly leading the witness to begin reasoning that number 4 is the one that police suspect. In this case, the automatic and deliberative processes compete with each other for the final decision. In other cases, the automatic and deliberative processes might corroborate each other. It is this relation between the automatic and deliberative processes that accounts for the name we have given to our conceptualization, namely the competition/corroboration conceptualization. Agreement (corroboration) between the automatic and deliberative processes should not only make the positive identification more likely, but also increase confidence in the identification. Disagreement (competition) between the automatic and deliberative processes should make positive identification less likely and should decrease confidence in any identification that occurs.

We argue that our competition/corroboration conceptualization of automaticity explains quite a bit of the available data. Accurate identification decisions are made faster than inaccurate identification decisions because stronger automatic responses require less corroboration. At the same time, even accurate identification decisions are made surprisingly slowly (e.g., 10–12 seconds or longer) in comparison with purely automatic decisions, because some deliberative processes also are involved. Responses to target-absent lineups are slow because no appreciable automatic process is involved; it is almost all a deliberative process. Furthermore, findings showing that the manipulated addition of deliberative processes serves to reduce accuracy are consistent with our idea that deliberative processes can compete with automatic processes in terms of the final decision made by the eyewitness. The addition of certain types of deliberative processes increases the chances that the eyewitness will become aware of some deliberative information that is inconsistent with the automatic processes and, assuming the automatic processes are usually more likely to be correct, this additional deliberation will reduce decisional accuracy. The latter idea suggests that there will be special circumstances in which reducing deliberative thoughts (e.g., by requiring quick decisions or using cognitive load
manipulations) will actually benefit the accuracy of eyewitness identifications, even though studies to date have not been able to demonstrate this effect. Specifically, when cues are present in the lineup task that suggest an answer that is inconsistent with the automatic response, then limiting deliberation will serve to increase reliance on the automatic process and thereby increase accuracy. Perhaps previous studies have failed to find that limiting deliberative processes increases identification accuracy because there were no cues present in the situation that would have led to inconsistencies (hence competition) between the automatic and deliberative processes.

The idea that automatic and deliberative processes do not necessarily operate exclusively is consistent with recent attempts to model eyewitness behavior mathematically (Clark, 2003). Clark’s WITNESS model suggests that identification decisions by eyewitnesses are based on a fairly even combination of absolute and relative match information. To the extent that absolute and relative judgments reflect automatic and deliberative judgments, respectively (which, as we have argued above, is probably not a perfect reflection, but one that is quite likely on average), his data would suggest a non-exclusive operation of these two types of cognitive processing, consistent with the work of Logan (1988) and with our competition/corroboration conceptualization.

It may be tempting, after all this discussion, to conclude that automatic processes are necessarily desirable and deliberative processes are necessarily undesirable with respect to eyewitness identifications. We would caution against this assumption. As previously discussed, a relative judgment strategy, which tends to lead to more false identifications than an absolute judgment strategy, may operate, in part, through an automatic process. Conversely, deliberative processing may at times be beneficial. Suppose, for instance, that an eyewitness viewed a photo lineup, made no identification, and then was shown a live lineup. At the live lineup, only one person (the suspect) was also a member of the photo lineup. The result might be an automatic reaction to the suspect based merely on the fact that he is the only one who evokes familiarity. Here, we would hope that the eyewitness might deliberatively reason that, because there is only one person in common between the photo lineup and the live lineup, that person might be familiar merely because that person was previously viewed in photos. If so, then the witness might use this deliberative reasoning to avoid an identification based on mere familiarity. Deliberative thinking may thus allow witnesses to override automatic judgments that are not based on memory of the criminal.

If automatic processes are not necessarily desirable and deliberative processes are not necessarily undesirable, what value is contained in the automatic/deliberative distinction? We believe that judgments high in automaticity are, in fact, desirable (but only to the extent that they reflect recognition memory of the criminal. To the extent that judgments high in automaticity reflect something other than recognition memory of the criminal, they can be deleterious. It is important, therefore, to empirically develop tools to differentiate between types of automatic processes. Additionally, understanding the root from which automatic and deliberative processes stem (which, as we have argued, may be the quality of memory that eyewitnesses have of the criminal) may focus researchers on the development of procedures that will maximize eyewitness accuracy. Finally, understanding what processes are more deliberative and what processes are more automatic
allows us to determine the extent to which witnesses can accurately report on their cognitive processes. Because people have low introspective awareness of automatic processes, the automatic/deliberative distinction allows us to learn when we can and when we cannot trust eyewitnesses’ self-reports.

**SUMMARY AND PROSPECTUS**

We have attempted to describe an applied theory of lineups. The theory is not one of how faces are remembered or how witnessing conditions affect encoding or yet another theory of memory processes. Instead, it is a theory of the lineup task, and it is deeply embedded in the ecology of forensic evidence. We described, for instance, how the function of a lineup is not to test the witness’s memory, but to test a hypothesis about whether the suspect is the perpetrator. Accordingly, the result of a lineup needs to speak to the status of the suspect given the response of the witness, not the response of the witness given the status of the suspect. This means that base rates regarding the status of the suspect (guilty or not) have a profound impact on outcomes. We also described a taxonomy of outcome distributions that distinguish sharply between false-positive identifications when the perpetrator is in the lineup and false-positive identifications of an innocent suspect when the perpetrator is not in the lineup. This distinction is required by the ecology of the lineup but is not part of any traditional theory of memory. Indeed, as we have discussed in this chapter, the lineup task involves considerations that are clearly extra-memorial, such as external influences from the lineup administrator. We also discussed the pleading effect, a factor that has largely been ignored in the eyewitness identification literature, but has profound implications for understanding how error rates at the level of the lineup might multiply in the subset of cases that go to trial. Finally, we have discussed broad psychological processes that can operate automatically, deliberatively, or both.

We stated that a good applied theory of lineups should illuminate important gaps in our knowledge about lineups. The gaps in our knowledge that would help further develop this applied theory are readily apparent. Here, we describe three important gaps. First, our analysis of outcome distributions proves that nonidentifications must have diagnostic exonerating value. But how do law enforcement and prosecutors treat nonidentifications in terms of revising their subjective probabilities that the suspect is the perpetrator? The suggestion that the legal system too readily dismisses nonidentification evidence was made over 20 years ago (Wells & Lindsay, 1980), and yet no progress has been made in finding out whether this is the case. Second, we have shown the profound role played by the perpetrator-present versus perpetrator-absent base rate variable, but we have made no progress in estimating this base rate in actual cases or in assessing practices of criminal investigators that drive this base rate. Third, we proposed a reanalysis of the distinction between automatic and deliberative processes, whereby a third variable—quality of memory for the criminal—drives both identification accuracy and decision process, and in which automatic processes may be best conceptualized as a continuous variable. Furthermore, we have suggested that deliberative processes can operate in parallel...
with automatic processes and can either corroborate or compete with automatic processes in terms of the final decision made by the eyewitness.

Most psychological theories confine themselves to psychological processes, and, in this sense, our sketch of an applied theory of lineups is unconventional. A conventional theorist in psychology might describe lineup identifications in terms of acquisition, storage, and retrieval processes of memory. As an applied theory, however, we posit that the relevant characteristics of lineup theory must also be steeped in an understanding of the structure of the task, the role of nonmemorial variables (such as social influence), the nature of the problem facing the trier of fact (e.g., assessing the status of the accused from the behavior of the witness, not vice versa), interactions with other types of pre-lineup evidence (which affect the base rate for the perpetrator being in the lineup), and an understanding of the role of automatic and deliberative processes (which relate to the ability of the witness to accurately testify about variables affecting his/her identification). Clearly, lineup theory must progress well beyond the sketch we have provided here, and we do not pretend that what we have written here is anything more than a beginning. Nevertheless, we believe that we have sketched some of the components that will need to be a part of an applied theory of lineups.

REFERENCES


