1

Person Descriptions as Eyewitness Evidence

Christian A. Meissner
University of Texas at El Paso

Siegfried L. Sporer
University of Giessen, Germany

Jonathan W. Schooler
University of British Columbia, Canada

Two teenage girls were enjoying their family vacation in a hotel hot tub one evening. Shortly after their parents had left them, the girls were approached by a stranger, who proceeded to join them in the hot tub. Following several minutes of conversation, the stranger attempted to molest the older girl by touching her "private parts." The older girl struck the stranger in the face and told him to stop, and instructed the younger girl to find their parents. After the older girl shouted "rape" several times, the stranger finally exited the hot tub, gathered his belongings, and ran from the area. The girls would later describe the stranger as a male in his twenties, with no shirt, wearing tan/brown shorts and a shell necklace. He had dark hair, and a dark/suntanned complexion. Upon receiving the description, detectives released a BOLO ("be on the lookout"), and a suspect matching the description was detained 45 minutes later as he walked on the beach about a half-mile from the hotel. The girls were brought to the suspect and together identified him as the stranger they had encountered. The suspect was arrested for the crime, but prosecutors would later drop the charges when the suspect provided a detailed (and corroborated) alibi for his whereabouts at the time of the incident. Simply put, detectives had detained the wrong person.

Person descriptions represent an important element for detectives in the investigation of any crime. Unfortunately, the descriptions provided by witnesses or victims tend to be rather nondistinct and, like the description provided by the teenage girls above, can frequently apply to many people in the vicinity of the crime. Although descriptions

are most often useful for locating a suspect in the immediate aftermath of an incident, they are also used throughout a criminal investigation to identify potential suspects from mug books, to construct sketches or composites of a suspect, and as a basis for selecting fillers when investigators are constructing a lineup identification parade and subsequently assess the "fairness" of that lineup. In addition, witness descriptions are regularly introduced at trial as a means for demonstrating the congruence between the suspect and a witness's memory. In *Neil v. Biggers* (1972), the U.S. Supreme Court indicated that witness descriptions could be used as one basis for determining the accuracy of a witness. However, as discussed in this chapter, the relationship between a witness's description and his or her ability to perceptually identify the actual perpetrator is not clear-cut.

Given the importance of person descriptions as eyewitness evidence, psychologists and criminologists have conducted a wealth of research aimed at establishing what is known about the content and veracity of person descriptions, as well as factors that may positively or negatively influence a witness's ability to provide an accurate description. The current chapter provides a review of this research, including a discussion of psychological factors that may influence person descriptions at encoding (e.g., alcohol, stress, illumination, distance, etc.), the effects of delay and repeated descriptions over time, the role of person variables (e.g., age, gender, race, etc.) and individual differences, and the influence of misinformation from investigators and/or co-witnesses. In addition, we address the variety of recall techniques that have been explored to improve the quality and quantity of person descriptors, and the relationship between such description procedures and witnesses' subsequent attempts at perceptual identification of a suspect (i.e., the verbal overshadowing effect, Schooler & Engstler-Schooler, 1990; or the use of person descriptions as retrieval cues, Cutler, Penrod, & Martens, 1987; Sporer, in press).

CONTENTS OF PERSON DESCRIPTIONS IN REAL CASES

Quantity and Quality of Descriptors

A number of archival studies have examined the content of person descriptions in real cases. Likely the most well-known study was conducted by Kuehn (1974). This study involved the analysis of person descriptions contained in 100 police protocols of cases of bodily injury, rape, and robbery in Seattle, Washington. Statements were taken from the witnesses immediately after the incident, and all perpetrators in the sample were strangers. Unfortunately, it is not clear from Kuehn's report whether the descriptions were rendered as free descriptions or were the result of some standardized questioning scheme employed by the local police. The number of details contained in the descriptions was fairly meager overall (with a maximum of nine descriptors)—on average there were 7.2 descriptors, whereas most witnesses reported 8 or 9 features. Only four victims were unable to provide any details at all. In descending order of frequency, gender, age, height, build, race, weight, complexion, and hair color were mentioned. With the exception of eye color (23%), all features were named by more than 70% of all vic-

tims. Kuehn concluded from these data that witnesses were able to convey a general impression about the perpetrator but could not provide more specific features, like hair or eye color.

In contrast to Kuehn (1974), Yuille and Cutshall (1986) examined a single shooting incident, involving a total of 21 witnesses, of which 13 collaborated in a follow-up research interview. Surprisingly, the witnesses' reports were remarkably elaborate and highly accurate, even after the 5-month delay between the incident and the research interviews. Based on these results, the authors questioned typical laboratory findings that have capitalized on witness errors since the heyday of eyewitness testimony research by Stern (1902) and Muensterberg (1908). However, one potential explanation for the findings of Yuille and Cutshall may be that the type of case, which apparently was quite spectacular, was likely to have involved multiple interviews of the witnesses, many conversations between witnesses and family/friends, as well as ample opportunity to read about the events in the local press. We speculate that these multiple information exposures may have served as opportunities for witnesses to rehearse these events and thus improve recall (for more on the role of verbal and visual rehearsal in eyewitness recall, see Read, Hammersley, Cross-Calvert, & McFadzen, 1989; Sporer, 1988, 1989).

Overall, Yuille and Cutshall (1986) found that the police interview had rendered a total of 392 action, 180 person description, and 78 object description details, many of which turned out to be correct (82%, 76%, and 89%, respectively). These figures were even higher for the research interview, which asked additional questions that were of primary interest to a memory researcher but not to a police investigator. Yet, despite the large number of correct details elicited by both the police and the researchers in the interviews there were also some errors in the person descriptions, most of which referred to estimates of height, weight, and age (23% errors out of a total of 46 such statistics in the police interview). Such estimates were deemed errors if they were outside of an acceptable range determined by the authors (i.e., plus or minus 2 inches or years, or 5 pounds). Second in errors were faulty descriptions of style and color of hair, as well as style and color of clothing (18%). Problems with descriptions of clothing, particularly memory for colors, were noted long ago by Muensterberg (1908). Cutshall and Yuille (1989) were subsequently able to elicit a greater number of details (although not a greater proportion of accurate details) than the police from witnesses of shootings and of bank robberies up to 2 years after the incident.

Sporer (1992a) analyzed criminal records containing 100 witnesses (46 male, 54 female) who provided a total of 139 person descriptions of perpetrators of capital crimes (mostly cases of robbery and rape). About half of the descriptions were from witnesses who were themselves victims or bystanders involved in the criminal action; the other half were from persons who had observed the criminal outside the context of the crime itself. Overall, person descriptions tended to be rather poor; the number of descriptive details mentioned ranged from 1 to 48 details ($M = 9.71$; $SD = 7.03$). Almost one-fourth of all descriptive details referred to general information such as height, age, and race, where height estimates frequently referenced some (unknown) population norm (e.g., "average height", "normal body figure," etc.). Another 31% of descriptors referred to clothes, and 30% described the face of the perpetrator. Some witnesses also men-

tioned jewelry or the dialect spoken. Close to 5% of the descriptors referenced personality characteristics (which are useless when investigators are trying to find a person to arrest, but may promote subsequent recognition of the person because of the deeper level of processing possibly involved at encoding; see Sporer, 1991). It is also noteworthy that quantity and pattern of descriptions found in this archival study closely resembled those of a staged event study in which a confederate had interrupted a lecture to take away a slide projector (Sporer, 1992b).

Of the facial descriptors analyzed by Sporer (1992a), the majority referred to the upper half of the face, particularly the hair of the perpetrator. This finding confirms earlier studies on contents of facial descriptions (Ellis, Shepherd, & Davies, 1980; Laughery, Duval, & Wogalter, 1986; Shepherd, Ellis, & Davies, 1977) and on the importance of upper portions of the face in the recognition process (e.g., Fisher & Cox, 1975). In fact, the cheek and chin (but also the forehead) were rarely mentioned in these descriptions. Although reference to hair (about 16% of all descriptors) appeared to be the most dominant single descriptor, it is also the most problematic and is likely to be of little help in the pursuit of a criminal because hair style can be most readily altered in comparison with other more permanent features (e.g., inner features of a face). Similarly, the large number of references to the clothing of the perpetrator is generally of limited value when police are attempting to locate a perpetrator.

Lindsay, Martin, and Webber (1994) examined the descriptions of 105 criminals published in the Kingston, Ontario newspaper (*The Whig Standard*) and compared their completeness with that of 100 descriptions (across five targets) obtained from a series of laboratory studies. Participant-witnesses viewing staged crimes were most likely to report clothing (99%), hair color (90%), and height (86%), whereas less than 50% reported such obvious descriptors as gender, age, or race/ethnicity. The most frequently reported feature of the face was the eyes (43%), and all other features were reported less than 25% of the time. Witnesses to real crimes were significantly more likely to report gender (96%), hair color (38%), clothing (60%), and race/ethnicity (25%), and facial features were provided in less than 10% of the sample. Although the results of Lindsay et al. indicated that laboratory witnesses provided more complete descriptions than real witnesses (7.35 vs. 3.94 features, respectively), they more generally concluded: "The data strongly support our concern that eyewitness descriptions are frequently vague" (p. 531).

Van Koppen and Lochun (1997) reported a large-scale archival analysis of person descriptions in 431 robbery cases. A total of 1313 witnesses provided 2299 descriptions of the offenders. Descriptors were subdivided into 24 permanent descriptors (e.g., gender, skin color) and 19 temporary characteristics (e.g., particulars of clothing, type of mask). Similar to Sporer's (1992a) findings, the completeness of the descriptions was rather poor. Of the possible maximum of 43 descriptors, the median number provided by each witness was 8 (interquartile range = 6). Permanent features were mentioned more frequently (median = 5, interquartile range = 5) than temporary characteristics (median = 2, interquartile range = 3). Considering that gender, appearance (including race), and skin color were among the most frequently mentioned permanent characteristics (characteristics that are likely the most obvious to any observer), the paucity of these descriptions becomes even more dramatic. Less than 5% of the descriptors referred

to inner features of a face (eye color, nose, face color or complexion, mouth, eye shape, teeth, earrings, chin, ear size, ears protruding), which are considered most important for identifying another person (Ellis, 1992). Of the temporary characteristics, the majority of descriptors referred to hats (51%) and hat color (31%), as well as jackets (28%), coats (25%), and trousers (26%), and their respective colors (28%, 22%, and 18%).

Van Koppen and Lochun's (1997) analysis was not restricted to the quantity of information recalled as in Sporer's (1992a) study, but also sought to analyze the accuracy of descriptions by validating the descriptions by witnesses against the descriptions contained in the police database used in the Netherlands. Although more elements of the descrptors were correct than incorrect, the majority of crucial facial descriptors were wrong (e.g., accuracy of facial descriptors included: eye color = 36%; nose = 35%; mouth = 39%; chin = 38%). Most strikingly, almost all descriptors of facial hair (beard and mustache) failed to match the police database. Given that perpetrators may have changed these aspects of their appearance over time, however, the latter finding is difficult to interpret. Interestingly, there was a negative correlation between accuracy and completeness, indicating that when witnesses did provide more extensive descriptions their accuracy suffered.

Estimates of Height and Weight

Almost all person descriptions contain references to the perceived height, weight, and age of the perpetrator (Kuehn, 1974; Sporer, 1992a; van Koppen & Lochun, 1997; Yuille & Cutshall, 1986); however, authors differ in their interpretation of existing data regarding the extent to which such estimates are accurate. Some authors have defined accuracy as estimates falling within a certain range of "true" values (e.g., true value plus or minus 2 inches or 5 pounds; see Yarmey & Yarmey, 1997; Yuille & Cutshall, 1986), concluding that estimates appear to be rather accurate. Then again, treating values with a difference of 4 inches (almost 10 cm) in height as "accurate" would allow an estimate of 170 cm to be equivalent to one of 180 cm, values that are substantially below or above the population average (see Flin & Shepherd, 1986; Sporer, 1996).

When the accuracy of estimates for height and weight are defined as the correlation between the actual values and their estimates, these correlations are well below their maximum possible value. For example, Janssen and Horowski (1980) reported that the average correlations between the actual and estimated heights in a series of studies with students aged 10 to 18 fluctuated between $.26 < rs < .90$. As might be expected, the correlations were smaller for younger children than for older teenagers. This age effect could be either a function of the restricted experience of the smaller children with numbers (see also Davies, 1996) or a result of the smaller children's own height, which seems to assist adults in gauging their estimates of another person.

Next to the target's true height and weight, probably the most important determinant of this type of estimate is the witness's own height and weight, perhaps modified by his or her knowledge (or better, supposition) of what the average population norm might be for a typical middle-aged male or female. Flin and Shepherd (1986) have presented a comprehensive and representative study on this topic. The authors had 588 participants

estimate the height and weight of male targets (using a total of 14 targets of differing heights and weights). Each target was accompanied by a second person, the context person, who asked the participant for directions. Thereafter, the context person returned to the participant and asked for an estimate of the target's height and weight, as well as the participant's own height and weight. Overall, Flin and Shepherd found evidence for an own-anchor effect in which participants used their own height or weight as a reference to judge that of the target person. In contrast, neither the context person's height nor his or her weight appeared to influence participants' estimates, as might have been expected if participants were to compare the two individuals side by side. Generally, participants underestimated the target person's height and weight. There was also a tendency for participants to underestimate the height of taller targets and to overestimate the height of shorter targets—a finding that could be interpreted as regression to the mean. Flin and Shepherd explained this finding with reference to subjects' knowledge of population norms, which might induce observers to shy away from extreme judgments. Hence, very tall or very heavy targets were more likely to be underestimated.

FACTORS THAT INFLUENCE DESCRIPTION ACCURACY/COMPLETENESS

Consistent with any memory task, the accuracy or completeness of person description is likely to be influenced by a host of factors, including those present at encoding (e.g., observation to view, anxiety or stress, etc.) or throughout the retention interval (e.g., length of the interval, post-event misinformation, etc.). The current section will review the available laboratory, field, and archival research on such factors, as well as witness or target variables (e.g., age, race/ethnicity, gender, etc.) or other individual difference variables that might influence description performance. Finally, various methods for obtaining person descriptions are discussed here for their influence on accuracy and completeness.

Encoding-Based Factors

Opportunity to View. It has been assumed that many if not most crimes happen at night; however, few studies have directly assessed the influence of illumination levels on person descriptions. From the perception literature we know that color vision is dramatically reduced at low levels of illumination, which implies that descriptions of clothing or hair color given under these conditions must be treated with caution. In addition, less information can be extracted under low levels of illumination (G. R. Loftus, 1985; Reinhardt-Rutland, 1986), which should lead to poorer descriptions. One study conducted by Yarmey (1986) has confirmed these extrapolations to person descriptions. More specifically, Yarmey examined eyewitness recall and identification of an event under conditions representing daylight, beginning of twilight, end of twilight, and night vision. His results indicated a significant influence of illumination level on witness recall, including details of the perpetrator, the victim, and the environment. As might have been expected, recall was superior during the daylight and beginning of twilight conditions.

Although it has been assumed that the opportunity to view a target person (i.e., distance between or duration of the event) should significantly influence the accuracy or completeness of person descriptions, only a handful of field studies have attempted to investigate such factors. For example, Yarmey, Jacob, and Porter (2002) conducted a study in which participants interacted with a target person for 5 seconds or 30 seconds and were subsequently asked to describe the encounter. As expected, their results indicated that person descriptions (particularly for clothing) were superior when participants had a longer time to observe the target person. Another aspect that appears to be important regards whether the witness encodes information about a perpetrator with the intent of later recalling it from memory. Along these lines, Yarmey (2004) found that instructions to intentionally encode information from the event for a subsequent memory test led to superior recall of person descriptions (again particularly for articles of clothing).

Although both laboratory and field research on such factors has been minimal, there are some archival analyses of criminal records that have explored the importance of viewing conditions. Despite claims for the superior ecological validity of archival studies (Yuille & Cutshall, 1986), the problem with archival analyses is that the accuracy of the descriptions generally cannot be determined—rather, a proxy for accuracy must be created with respect to the precision of the description, or its relative consistency with that of the individual found guilty for the crime. In Sporer's (1992a) study, the mean number of descriptors, length, and precision of person descriptions were coded and related to low, medium, and high levels of a host of potentially relevant factors, including illumination, duration of event, and time to observe. The categories "low," "medium," and "high" are not to be taken literally, as they may take on different meanings with respect to the particular variable coded (e.g., "high illumination" was operationalized as bright daylight or good artificial lighting). Level of illumination had the expected effect (such that greater illumination led to more complete person descriptions), whereas duration of the incident and time estimated for the target to be in view did not seem to influence description completeness. Similarly, van Koppen and Lochun (1997) found that better illumination and shorter distances between the witness and perpetrator were associated with greater frequency of person descriptors. Whereas both of these studies supported the predicted linear relationship between opportunity to view and recall, Kuehn's (1974) archival analysis found worse performance for twilight conditions than for observations either at bright daylight or at night.

Stress or Anxiety. Eyewitness events are generally considered anxiety-provoking situations in which the victim or witness is likely to experience a great deal of stress during the encoding process. Consistent with this notion, a number of studies of eyewitnesses have suggested that high levels of stress or anxiety impair memory by restricting attentional and executive processes at encoding and thereby prevent the consolidation of information into a coherent event sequence (see Deffenbacher, 1983, 1994). On the other hand, other studies suggest that stress may increase participants' memory for central details (Christianson, 1992) and that the negative effects of stress (at least in some cases) may reverse with the passage of time (Burke, Heuer, & Reisberg, 1992; Christianson, 1984; for a general review see Deffenbacher, Bornstein, Penrod, & McGorty, 2004;

Schooler & Eich, 2000.) With regard to person descriptions, several laboratory studies have demonstrated impairment in accuracy and completeness as a result of stress or anxiety. For example, Clifford and Hollin (1981) varied the violence of a to-be-remembered event and found that participants in the violent conditions were less likely to recall details of the perpetrator (see also Loftus & Burns, 1982). In their recent meta-analysis on the topic, Deffenbacher et al. (20004) found that heightened anxiety led to significant decrements in recall accuracy (Cohen's $d = -.31$) across studies.

The presence of a weapon, which may be accompanied by stress or fear, has also been shown to divert a witness's attention away from the face of the offender. A number of studies have investigated the possibility that the presence of a weapon is associated with impaired recall of details of the perpetrator or event. Consistent with the aforementioned research, studies of the "weapon focus" effect have generally demonstrated a significant influence of the presence of a weapon on person description accuracy (see meta-analysis by Steblay, 1992). Recent research by Pickel (1998, 1999) has indicated that the unusual or unexpected nature of a weapon may be responsible for the observed effect on description accuracy, when contrasted with the "threat" posed by the object.

Archival studies of eyewitness testimony have also attempted to assess the influence of anxiety, stress, or the presence of a weapon on the accuracy or completeness of person descriptions. Given that stress in criminal situations could not be observed (or manipulated) directly, the amounts of anxiety and arousal were coded retrospectively by classification of an event on the basis of the reports emerging from police records (e.g., presence of a deadly weapon, bodily injury, etc.) or of self-reports of anxiety provided by the witnesses in the course of testimony. In Sporer's (1992a) study, three groups of witnesses were compared: victims, bystanders participating in the event without being victims, and other witnesses who were questioned by the police about the perpetrator during the investigation but were not themselves directly involved in the case (e.g., the owner of a gunshop where the perpetrator bought his weapon). Overall, the most striking finding of this analysis was that none of the various ways in which stress had been coded seemed to indicate the expected deterioration in witness recall for high levels of stress and its associated variables. In fact, there even appeared to be a (linear) increase in descriptive details as a function of some of these stress-related variables (e.g., greater reported anxiety was associated with a greater number of details). An analysis of stress conducted by Yuille and Cutshall (1986) showed similar results, whereas an analysis conducted by van Koppen and Lochun (1997) demonstrated results consistent with the laboratory and field research reported earlier (i.e., high levels of stress associated with impaired recall performance). A more recent archival study by Wagstaff et al. (2003) demonstrated null effects on the accuracy or completeness of person descriptions. The general inconsistency observed between laboratory or field research and archival research may potentially be accounted for by length of the retention interval. Laboratory studies have typically used short retention intervals that are known to sometimes give an advantage to nonstressful memories, whereas archival studies typically involve longer retention intervals, which sometimes afford advantages to more stressful memories (Kleinsmith & Kaplan, 1963, 1964). It also possible that stressful experiences may be more likely to incur rehearsal, which could increase the amount of details recalled. Importantly, in none of these

archival studies was it possible to ascertain the accuracy of the descriptions, so all conclusions from the archival data must be drawn with caution.

Alcohol or Drugs. The consumption of alcohol or drugs is frequently associated with criminal activity (Yuille, 1986). Laboratory research has consistently demonstrated that alcohol consumption inhibits the encoding process when administered beforehand and thereby impairs subsequent recall of information (for a review see Sayette, 1999). However, research has been somewhat limited in examining the influence of alcohol or drug usage on the accuracy or completeness of eyewitness descriptions. One of the few empirical studies examining the effect of alcohol consumption on witness recall was conducted by Yuille and Tollestrup (1990). In general, the authors found that consumption of alcohol significantly impaired participants' ability to recall details (in both frequency and accuracy of recall) of the event and/or target person, regardless of whether the participant recalled immediately (and under the continued influence of alcohol) or 1 week later. Read, Yuille, and Tollestrup (1992; Experiment 1) subsequently found similar effects. In his archival analysis, Sporer (1992a) also found that when witnesses had consumed alcohol they were less able to report details about the perpetrator's appearance.

More recently, Yuille and his colleagues (Yuille, Tollestrup, Marxsen, Porter, & Herve, 1998) investigated the effects of marijuana use on eyewitness memory. Prior research had generally shown detrimental effects of marijuana on memory recall (cf. Murray, 1986). The results of Yuille et al. demonstrated that marijuana use significantly impaired the completeness of witnesses' recall regarding the event or target person. This effect, however, was moderated by the timing of recall such that the impairment of recall associated with marijuana use was present only when participants were questioned immediately after the event. When participants in the marijuana and control conditions were questioned after a 1-week delay, no differences in completeness of recall were observed. In contrast to completeness of recall, no significant effects of marijuana use were found when accuracy of recall was considered. It is apparent that further research is needed to evaluate the influence of alcohol and drugs on eyewitness recall.

Retention Factors

To the layperson it may sound like a truism that accurate retrieval of information should deteriorate following increased levels of delay; however, the form of the postulated forgetting function varies with the type of material (e.g., visual vs. verbal) as well as the form of the memory test (e.g., recall vs. recognition; see Shepherd, Ellis, & Davies, 1982; Shepherd, 1983; Sporer, 1989; Wixted & Ebbesen, 1997). The current section discusses the available research (both laboratory, field, and archival) regarding the influence of retention factors on person descriptions, including the length of the delay, the strength of the memory trace, and the intrusion of post-event information.

In general, laboratory research has shown significant detrimental effects of delay in the accuracy and completeness of person descriptions. For example, Ellis, Shepherd, and Davies (1980) had participants describe one face immediately after viewing it, and

another either 1 hour, the next day, or 1 week following exposure. Participants remembered significantly fewer details after 1 week compared with the two shorter retention intervals, and memory loss was rather equally distributed across specific facial features. The accuracy of person descriptions also declined significantly with the longer delay interval. In a similar laboratory experiment, Meissner (2002) found significant losses in both the completeness and the accuracy of facial descriptors when participants provided a description either immediately or following a 1-week delay.

In their archival analysis, van Koppen and Lochun (1997) observed a pattern consistent with the aforementioned laboratory studies, such that witnesses provided fewer person descriptors following longer retention intervals. In contrast to this study, Yuille and Cutshall (1986) and Cutshall and Yuille (1989) emphasized strikingly high levels of recall from witnesses of real crimes as late as 2 years after the incidents. As mentioned previously, these high levels of performance were likely mediated by repeated questioning (and rehearsal) prior to recall at the time of the study (see Sporer, 1989).

It should be noted that the course of time alone is unlikely to have a detrimental effect on recall; rather, both the strength of the initial memory trace and interference from a variety of activities during the delay interval are likely the major influence of a witness's ultimate recall of person descriptors. Generally referred to as "post-event information," witnesses may obtain information during the retention interval (either deliberately or unintentionally) through a number of sources or tasks that they engage in. For example, overhearing a description provided by another person or being shown an erroneous facial composite or sketch can lead the witness to incorporate erroneous details into his or her own description of the perpetrator, and the likelihood of such post-event information influencing subsequent recall has been shown to increase following a long retention interval (Loftus & Greene, 1980; Loftus & Ketcham, 1983; Shaw, Garven, & Wood, 1997; Sporer, 1996b). The related effects of misleading questioning by investigators (referred to as "misinformation") and collaborative recall with another witness (or "co-witness" effects) are discussed below.

Witness and Target-Person Variables

As in the eyewitness identification literature, a number of witness and target variables (e.g., gender, age, ethnicity, etc.) appear to influence the accuracy and completeness of person descriptions. This section reviews the available literature on such variables.

Gender. Although many studies on eyewitness memory have included both male and female participants, few have analyzed gender differences. Several studies conducted by Yarmey (1986, 1993, 2004) have generally indicated few differences in the recall of men and women. When differences were noted, they typically involved responses to specific attributes that women may have been more likely to attend to at encoding (e.g., jewelry, hair color or length, and weight; see Yarmey, 2004), or they involved more complex interactions between variables (such as levels of illumination; see Yarmey, 1986). In several studies, Yarmey noted that men appeared more confident in their responses than women (Yarmey, 1986, 1993).

MacLeod and Shepherd (1986) have drawn attention to gender differences in an archival study of criminal assault cases. Similar to research by Yarmey (1986, 2004), gender differences were found to covary in a complex manner with such variables as the type of questions analyzed (e.g., action details vs. descriptive details; statements referring to self, victim, accused, or periphery) and the type of incident (involving injury of the victim or not). In his archival analysis, Sporer (1992a) reported that male witnesses provided on average longer descriptions than females (M = 7.50 vs. 7.10 number of lines in the protocol, respectively). In contrast, the number of descriptors and rated precision of statements showed an opposite but nonsignificant trend favoring females. Thus, it appears that although females may have said less quantitatively, they did not necessarily convey less information.

Child Witnesses. Although some studies have found that the relative accuracy in reports of children may not differ from that of adults (Goodman & Reed, 1986; Leippe, Romanczyk, & Manion, 1991; Marin, Holmes, Guth, & Kovac, 1979), adults' statements are likely to be much longer and more detailed than those of children (Davies, Tarrant, & Flin, 1989; Dent & Stephenson, 1979; Leippe et al., 1979; Marin et al., 1979). In contrast, a recent study conducted by Pozzulo and Warren (2003, Experiment 1) observed both greater accuracy and completeness of person descriptions provided by adults versus youths (ages 10 to 14). Further analyses indicated that adults were more likely to report features of the face, aspects of the body (i.e., height, weight, and build), and race of the perpetrator, whereas youths were more likely to report various accessories (e.g., belt or glasses). With regard to accuracy, youths were significantly less accurate than adults in describing interior facial features (e.g., eyes, nose, or mouth), aspects of the body, and the age of the perpetrator. In a follow-up study using a live event, Pozzulo and Warren (Experiment 2) observed the more classic pattern involving a greater frequency of person descriptors by adults when compared with youths, but no differences in the overall accuracy of features reported. The analysis of specific features was largely consistent with the first study, except that aspects of the target's clothing were more likely to be reported by adults in the sample. Recent research by Lindholm (2005) has also suggested that witnesses, particularly children and young adults, may actually perform better when recalling descriptions of target persons matching their own age group. Such own-age effects (similar to the cross-race effects discussed below) may result from a variety of experiential or motivational factors (cf. Sporer, 2001a), and further research on this topic seems warranted.

Saywitz (1995) has suggested that it may be important to adapt one's language when interviewing children such that questions are more comprehensible to young children. In particular, interviewers should use short sentences with a simple grammatical structure, common phrases, and proper names. They should avoid the passive voice, double negatives, and indirect questions. Before estimates are obtained, interviewers should also make sure that children understand concepts like size, distance, weight, age, and time, as well as particular body parts and various color names. For example, Dent (1982) reported large inaccuracies in estimates with children between 8 and 13 years of age. Furthermore, age estimates may suffer from children's lack of knowledge of facial cues to

aging (Ellis, 1992). Providing children (and adults) with possible ranges or specific anchors (Dent, 1982; Sporer, 1996b) or a color plate or color wheel may lead to better results for some aspects of person descriptions than free descriptions. In contrast, specific questions (e.g., *What was the color of her hair?*) may lead not only to more information but also to more inaccurate information than general questions (e.g., *What was her appearance?*; Dent & Stephenson, 1979).

One challenge in understanding the influence of misinformation on children relative to adults is that children may be both more likely to forget details of the original experience (including the appearance of the perpetrator) and more likely to forget any misinformation they receive about the individual after the fact (Schooler, 1998; Schooler & Loftus, 1993). Thus, it is possible that testing children following a delay (when they have had the opportunity to forget the misinformation) may provide the best opportunity for achieving veridical recall. In their classic review of the topic, Ceci and Bruck (1993) also posited that certain cognitive (e.g., memory trace strength or source-monitoring ability) and social (e.g., conformity to an authority figure) activities can mediate a child's susceptibility to suggestion in recall. Although the authors caution against the perils of suggestive questioning, they warn against completely discounting children's recall. In their own words, "children are able to encode and retrieve large amounts of information, especially when it is personally experienced and highly meaningful" (p. 434).

Elderly Witnesses. Aging in late adulthood has been shown to affect both the perceptual and memory abilities of witnesses (for a review, see Yarmey, 1996). Elderly witnesses (i.e., above 65 years), for example, are increasingly more likely to demonstrate deficits in their visual acuity at night, and in their ability to perceive depth and to adapt to darkness. Both color vision (particularly blue and blue-green) and memory for colors are also likely to decline with age. At later age levels, individuals are also more likely to demonstrate difficulty with source monitoring (Cohen & Faulkner, 1989; Henkel, Johnson, & de Leonardis, 1998), which may be crucial in many eyewitness situations (e.g., to counter the influence of suggestive questioning).

In a field experiment in which both showup and lineup identifications were administered to participants ranging in age between 18 and 65 (Yarmey, Yarmey, & Yarmey, 1994), 651 individuals were randomly approached in public places and asked for directions by one of two young adult, female confederates. The duration of exposure to the target was approximately 15 seconds. Two minutes later the witness was approached by a female investigator and was asked to describe the target and to identify her face and voice. With regard to description accuracy, young adults (18–29 years of age) were significantly superior ($M = 72\%$) to middle-aged witnesses (30–44 years of age) ($M = 61\%$), who in turn were superior to older adults (45–65 years of age) ($M = 54\%$). These results comport with prior research conducted by Yarmey and his colleagues (e.g., Yarmey & Kent, 1980), which indicated that "young adults on average were twice as complete and 20% more accurate in free narration in their descriptions of a criminal incident than were the elderly" (Yarmey, 1996, p. 268). Recent research by Searcy, Bartlett, Memon, and Swanson (2001) has demonstrated similar effects on person description completeness

and accuracy for young versus elderly adults. However, to the extent that most of these studies have used only young adults as targets to be observed and described, these studies may reflect as much an in-group bias in the form of an own-age effect (Sporer, 2001a) as deficits in the memory of elderly witnesses.

Cross-ethnic Differences. Although more than 60 studies have investigated recognition memory for own- versus other-race faces (for reviews see Chance & Goldstein, 1996; Meissner & Brigham, 2001b; Sporer, 2001a), very few studies have attempted to determine whether participants differ in the way they describe faces of their own and another race (Sporer, 2001b). Those that have investigated descriptions of own- versus other-race faces have suggested that individuals attend to features deemed relevant to own-race faces and further attempt to apply this encoding scheme inappropriately when examining other-race faces (Ellis, Deregowski, & Shepherd, 1975; Shepherd & Deregowski, 1981). For example, Ellis and colleagues (1975) demonstrated several differences in the type of features that black and white participants recalled (regardless of the race of face). Although Ellis and colleagues did not assess descriptions for accuracy or discriminability, they did note that white participants often reported rather "redundant" descriptions of black faces (e.g., "he has black skin, black, kinky hair and brown eyes") that would likely be indiscriminant upon later assessment (p. 123).

Fallshore and Schooler (1995) compared Caucasian undergraduates' ability to identify and describe African American and Caucasian faces. As is typically found, they observed the cross-race effect for lineup identification decisions, such that participants were better able to recognize Caucasian relative to African American faces. However, when description accuracy was assessed with the use of a communication accuracy paradigm in which subject-judges attempted to identify the faces based on witnesses' verbal descriptions, no cross-race effect was observed (although a numerical advantage was shown for the identification of other-race faces). Fallshore and Schooler speculated that differences in the pattern of results associated with cross-racial face recognition versus face description may be due to differential reliance on configural versus featural processing for own versus other race faces, respectively (see Rhodes, Brake, Taylor, & Tan, 1989). Accordingly, if the source of the own-race face recognition advantage were an enhanced ability to rely on configural information (Sporer, 2001a), then it follows that verbal description ability, which typically relies on featural knowledge (see Farah, Wilson, Drain, & Tanaka, 1998; Wells & Turtle, 1987), should not reveal such differences. Thus, although the relative dearth of studies on the topic clearly suggests the need for additional research, the absence of evidence for an own-race advantage for person description may reflect fundamental differences in the processes associated with face recognition versus description.

Methods for Obtaining Person Descriptions

Several methods of eliciting a person description have been developed over the years, from standard free recall approaches to feature checklists and techniques based upon

principles of cognitive psychology (e.g., the cognitive interview). In this section, we discuss research on the generation of person descriptions and their positive and negative effects. Along the way, we also address the role of leading questions and attempts at permitting witnesses to collaborate in generating a description, and we consider the effect of repeated questioning on the accuracy and completeness of person descriptions.

Free Recall vs. Leading Questions. Likely the most common technique used by investigators to obtain a person description involves a request for the witness to simply recall what he or she remembers about the perpetrator of the crime. Although such free recall descriptions are often quite accurate, unfortunately they rarely satisfy investigators, because of their likelihood of being incomplete with regard to critical details (Lipton, 1977). Thus, investigators will frequently follow up with more specific, close-ended questions to complete the description (e.g., *Do you remember the color of the man's hair?*). In addition, investigators may have previously received information regarding the perpetrator and so will attempt to confirm this information by inquiring about more specific details (e.g., *Did the man have red hair with long sideburns?*) or may include this information in the context of inquiring about another detail (e.g., *This man with the red hair and long sideburns, did he have any facial hair?*). Unfortunately, such leading questions can have rather harmful consequences for the witness's attempts at subsequent recall, as studies indicate that witnesses are quite likely to incorporate potentially inaccurate information ("misinformation") into their person descriptions (Loftus, 1975, 1979; Loftus & Zanni, 1975). For example, Loftus and Greene (1980) observed that participants who viewed a face and then heard a description of the face that was attributed to another witness later incorporated the verbal expressions of that witness into their description, even when the description was in error.

Feature Checklists. As noted above, one primary drawback to the use of free recall tasks regards the incompleteness of person descriptions. Witnesses will often vary in their output criterion for recalling details of an event (Koriat & Goldsmith, 1996), and a common difficulty with person descriptions involves the limited vocabulary that individuals have for describing the human face. In an attempt to alleviate this problem, researchers have sought to develop feature checklists that might aid witnesses in providing more complete (and useful) descriptions of the perpetrator they viewed. For example, Shepherd (1986; see also Shepherd & Ellis, 1996; Sporer, in press) and his colleagues have developed the Aberdeen Face Rating Schedule, which consists of some 50 items on which witnesses are asked to rate individual features of a face on five-point scales (for a published version of these scales, see Sporer, in press). Using these forms, observers are prompted to use certain features that otherwise they might omit or forget. However, accuracy of these descriptions might be poor, as people may frequently mark the middle ("normal") value of the scale when they either don't remember or guess the information (Sporer, in press). Nonetheless, forms of this type are useful both for communicating information to other agencies and for conducting computerized searches to identify individuals in mug shot databases who might be presented to the witness (cf. Pryke, Lindsay, & Pozzulo, 2000). A prototype of such a system was developed by psychologists at the

University of Aberdeen (see Shepherd, 1986; Shepherd & Ellis, 1996), and another similar system, SIGMA-IRIS, was used by the Austrian police (Zima & Zeiner, 1982).

One potential problem with the use of feature checklists regards their presentation of a rather exhaustive list of person descriptors, many of which the witness may never have attended to at encoding. The information elicited is either not very informative (as when witnesses mark default, "normal" values) or even incorrect, and the accuracy of the information is not related to the accuracy of a later identification (Sporer, in press). When the witness signifies the recollection of several features that are incongruent with the actual perpetrator, this may cause interference quite similar to the misinformation effects discussed above. In several studies, Wogalter (1991, 1996) has shown that such feature checklists (in contrast to a free recall or imaging task) can produce more incorrect features and subsequently interfere with witnesses' ability to identify the perpetrator. As a result, feature checklists may not provide the best means for collecting eyewitness information.

Collaborative Recall. Should witnesses be permitted to discuss their memory for the event with one another in generating a common, agreed-upon description for the perpetrator? There are, of course, potential benefits from collaborative recall, but there would also be potential costs of cross-contamination if witnesses were to share erroneous information with one another. Psychologists have studied this problem in the context of person descriptions, attempting to understand any benefits of permitting collaborative recall on the accuracy and completeness of descriptions, and the extent to which witnesses may adopt erroneous information provided by another witness into their descriptions. For example, Warnick and Sanders (1980) investigated the influence of group discussion of a previously viewed event on individual witness's subsequent recall. Their results indicated superior accuracy and completeness of recall for participants who had discussed the event in a group when compared with participants who recalled the information independently. Yarmey and Morris (1998) conducted a similar study, but had some participants also provide a consensus description of the perpetrator and event (some immediately, others following a 1-week delay). The results of Yarmey and Morris also indicated that group discussions led to more correct details being recalled when compared with individual attempts at recall, but no similar increase in erroneous details.

Given that witnesses' person descriptions are generally quite accurate, it seems reasonable that collaborative recall would have some positive effects on the amount of information recalled. But are witnesses particularly susceptible to adopting erroneous details that might be provided by another witness? To explore such a "conformity effect," Gabbert, Memon, and Allan (2003) created a situation in which witnesses viewed events differing in several key features. Witnesses were then later asked to discuss the event with another witness before providing a description independently. Consistent with previous studies of the misinformation paradigm (cf. Shaw et al., 1997), a rather substantial percentage of participants (71%) incorporated erroneous details provided to them by the co-witness. Thus, to the extent that a co-witness might provide erroneous information, collaborative recall may contaminate the person descriptions of others who participate in the discussion.

Repeated Questioning. Witnesses may be asked to provide a description of the perpetrator and event on multiple occasions, including immediately following the event, throughout the investigative process, in depositions and pretrial hearings, and finally (but most importantly) on the witness stand before a jury. To what extent might repeated questioning influence the veridicality of the information provided by the witness? The general cognitive literature has shown both positive and negative effects of repeated recall (Brown, 1923). For example, individuals may benefit from repeated attempts by recalling information or items that had not previously been reported (Payne, 1987; Roediger & Challis, 1989). To avoid confusion, we adopt the distinction between *hypermnesia*, that is, an increase in net recall (number of new details minus number of items lost), and *reminiscence*, that is, the gross recall of details provided at least once across a number of trials (Payne, 1987; Turtle & Yuille, 1994). In their study of eyewitnesses, Scrivner and Safer (1988) demonstrated hypermnesia during the repeated recall of event and perpetrator details from a previously viewed crime. Turtle and Yuille (1994) partially replicated these findings with longer retention intervals between successive recall episodes, demonstrating reminiscence but not hypermnesia. Bornstein, Liebel, and Scarberry (1998) further demonstrated that repeated testing can improve recall for details of a negatively arousing event.

In addition to the possibility of more complete descriptions, repeated testing has also been shown to preserve an individual's memory by strengthening associations that are retrieved (see Bjork, 1988). One important moderator, however, regards the retention interval prior to the first attempt at retrieval—to the extent that the retention interval is brief, more information may be preserved by the act of retrieval (Bahrick, 2000; Ebbesen & Rienick, 1998; Shaw, Bjork, & Handall, 1995). Ebbesen and Rienick (1998) varied the interval between exposure to a target and the first recall attempt (1 day, 7 days, or 28 days), and all participants provided a second recall attempt after 4 weeks. Their results indicated that, across all conditions, participants recalled about one less descriptor at the 4-week test ($M = 8.50$) than at all other tests ($M = 9.50$). Although the authors stress the fact that there was virtually no decline in the recall of personal attributes once a recall attempt was made, the percentage of errors for facial features, clothing color, and clothing style was still substantial. Even recall for the ethnicity of the person who participants had interacted with showed error rates between 13% and 23%. Nonetheless, these results do appear to demonstrate the predicted protection of person description memory afforded by repeated questioning. In a similar fashion, Dunning and Stern (1992) reported two experiments in which participants showed a (nonsignificant) tendency to recall more person information correctly, with no change in incorrect or confabulated details, over repeated reports. The interval between reports, however, was only 5 minutes, which is functionally quite different from the situation in which witnesses are repeatedly asked about events at different occasions separated by days or even months (Sporer, 1992a; van Koppen & Lochun, 1997).

In contrast to the benefits of increased completeness and maintenance of the memory, Roediger and his colleagues have demonstrated that repeated testing can also have rather paradoxical effects in which erroneous information may be reported and incorporated into subsequent recall episodes (see Roediger, McDermott, & Goff, 1997; Roediger, Wheeler, & Rajaram, 1993). For example, a study by Roediger, Jacoby, and McDer-

mott (1996) demonstrated that when participants were encouraged to recall erroneous information from a previously viewed crime, they were more likely to report that information in later attempts at recall (cf. Schooler, Foster, & Loftus, 1987). Meissner (2002) subsequently replicated the pervasive effects of self-generated misinformation in the context of person descriptions, particularly when participants were forced to report descriptors that they were unsure of.

Cognitive Interview. Over the years, researchers have been interested in devising techniques that might improve the accuracy and completeness of information obtained from witnesses. Likely the most well-known technique is the cognitive interview, which was initially developed by Geiselman and Fisher in the early 1980s (for a review, see Fisher & Geiselman, 1992). Overall, the cognitive interview consists of four main components: (1) context reinstatement, which includes mentally reinstating the environmental and personal context of the original event; (2) instruction to "report all" information, including partial information, even if it seems unimportant; (3) recounting the event in a variety of temporal orders; and (4) reporting the events from a variety of perspectives. With the use of the cognitive interview, a host of studies have shown that descriptions of persons, objects, and events can be reliably improved when compared with other standard (free recall) interview techniques.

In the first of these studies, Geiselman et al. (1984) obtained 11.00 correct details in response to open-ended questions about characteristics of a person from witnesses instructed with the cognitive interview, compared with 7.38 details by witnesses in a standard interview condition. Importantly, the cognitive interview did not lead to an increase in incorrect details. Whereas this basic pattern of results has been confirmed in studies with real witnesses (Fisher, Geiselman, & Amador, 1989), other studies have noted an increase in the recall of incorrect details gathered with the cognitive interview. For example, a study by Finger and Pezdek (1999) found that the cognitive interview increased the recall of both correct and incorrect facial descriptors when compared with a standard interview procedure. Confirming this pattern, a recent meta-analysis of 42 studies by Koehnken, Milne, Memon, and Bull (1999) revealed a large increase in the number of correct details elicited by the cognitive interview and a smaller, yet significant, increase in the number of incorrect details elicited. Furthermore, the meta-analysis indicated that accuracy rates elicited with the cognitive interview were about the same as accuracy rates achieved with traditional interview methods (84% vs. 82%, respectively). It should be noted that the majority of studies examining the cognitive interview have not focused on obtaining person descriptions per se, so further research in this direction seems worthwhile.

THE RELATIONSHIP BETWEEN PERSON DESCRIPTIONS AND EYEWITNESS IDENTIFICATION

So far our discussion has focused on the nature and quality of person descriptions. An important related issue involves the relationship between the description and identification of faces. This in turn leads to two distinct (albeit related) questions. First, what is

the relationship between the quality with which a witness describes a face and the accuracy with which he or she subsequently identifies it? Second, what is the influence of describing a face on its subsequent identification? As will be seen, the answers to both of these questions are not as intuitive as one might expect.

The Description-Identification Relationship

It seems quite reasonable that witnesses who are better at describing a perpetrator should also be better at identifying him. The intuitive nature of this relationship is inherent in the arguments in many eyewitness cases where inconsistencies between a witness's initial description of a perpetrator and the appearance of the suspect are highlighted to undermine the credibility of the identification. Both the U.S. Supreme Court (*Neil v. Biggers*, 1972) and the German Supreme Court have used the quality of person descriptions as indicators to evaluate the accuracy of person identifications in criminal trials (see Sporer & Cutler, 2003). Despite the appeal of the belief that a strong relationship should exist between face description quality and identification accuracy, research reveals that this relationship is at best very weak and often nonexistent. Although Sporer (1992b) reported a significant positive relation (assessed by a point-biserial correlation) between the number of descriptors and identification accuracy in a staged event study ($r = .28$), other studies have not confirmed this finding when focusing on the accuracy of descriptions. For example, Grass and Sporer (1991) staged another event in a classroom and then 1 week later had participants describe the target's appearance and respond to prompted questions about the target's appearance. Participants were then presented live simultaneous, live sequential, or photographic sequential lineups. These authors found no relationship between two judges' assessments of either the completeness ($r = -.06$) or accuracy ($r = -.04$) of the descriptions and identification performance. Similar failures to find a relationship between face description quality and recognition performance have been observed in a number of other studies (Pigott & Brigham, 1985; Sporer, in press). Furthermore, Wells and Leippe (1981) actually found a nonsignificant, yet sizable, negative relationship ($r = -.41$) between the accuracy of witnesses verbal recall of other aspects of the scene of a simulated crime and their identification of the target individual.

Although an absence of a relationship between person description quality and identification performance is by far the most common result, there are a few circumstances under which a relationship has been observed. Using the communication accuracy paradigm, Fallshore and Schooler (1995) examined the relationship between a description's quality and the ability of another individual to use a given description to identify the individual described from among a set of distractors. In the context of describing and identifying own- versus other-race faces, the authors found no significant relationship between description accuracy and identification performance for own-race faces ($r = .12$), but a significant relationship in performance on other-race faces ($r = .36$). This finding further supports the view that other-race faces may be recognized in a more featural manner than own-race faces (Rhodes et al., 1989; Sporer, 2001a). Accordingly, inasmuch as the recognition of other-race faces depends on the quality of witnesses' memory

of individual features, the veracity of the witnesses' memory for those features (as revealed by the quality of their descriptions) becomes predictive of their recognition performance. This finding also potentially offers a key for understanding why face description quality bears so little relationship to identification performance with own-race faces—namely, the two tasks may draw on fundamentally different types of knowledge, with the former depending on participants' memory for distinctive features and the latter depending on their nonverbal knowledge of the face in its entirety (see Farah et al., 1998; Wells & Turtle, 1987).

A second exception to the typical absence of a relationship between description quality and face recognition quality comes from studies that have compared the relative ease with which different faces can be described versus recognized. Wells (1985) showed participants multiple faces and then examined their ability to both describe and recognize each face. He found that distinctive faces tended to be easier to describe and to recognize than less distinct faces, thereby leading to a modest relationship between recognition accuracy and description quality ($r = .27$) across faces. Although this modest correlation does suggest that certain distinctive faces can be recognized on the basis of individual features, it certainly does not undermine the more common conclusion that typically little relationship between verbal description quality and recognition accuracy can be expected.

A final exception to the absence of a relationship between description quality and recognition performance has been observed in studies in which participants were forced to generate rather elaborate descriptions of faces and were later asked to identify these individuals in a lineup identification task (cf. Meissner, Brigham, & Kelley, 2001). In these studies, it appears that the elicitation of elaborate verbal descriptions may lead participants to generate inaccurate details, which then impairs their recognition performance. Indeed, several studies using such a paradigm (Finger & Pezdek, 1999; Meissner, 2002; Meissner et al., 2001) have found that incorrect details reported in participants' descriptions are predictive of subsequent identification errors.

In short, it seems that despite the clear intuition that witnesses who are better at describing a target should also be better at recognizing it, this relationship has proved to be quite elusive and generally weak. Though the absence of such a relationship may undermine this frequently relied-upon method for assessing the credibility of witnesses, it also provides an important link in our understanding of the nature of person descriptions—namely, that person descriptions may draw upon knowledge or cognitive processes that are very different from those invoked in the identification of a face. More specifically, person descriptions appear to encourage a focus upon verbalizable features of the face that are not always useful for perceptually individuating a given face from among similar distractors. In contrast, recognition of faces has been shown to involve a configural process in which features combine to create a nonverbalizable perceptual set that is stored and later accessed for pattern recognition (Farah et al., 1998). The exceptions to the incompatibility of these processes appear to involve faces that are recognizable based upon a distinctive local feature, or conditions in which retrieval of a face description distorts the veracity of the memory trace and interferes with subsequent identification.

The Influence of Person Descriptions on Identification: Verbal Overshadowing

The fundamental difference between describing a face and recognizing it also contributes to some counterintuitive findings regarding the influence of verbally describing a face on subsequent recognition of that face. Intuitively we might expect that describing a face would be helpful for subsequent memory performance, because it constitutes a form of verbal rehearsal, and verbal rehearsal is well known to enhance memory performance (e.g., Darley & Glass, 1975; Glenberg & Adams, 1978; see Sporer, 1989). There is some evidence that visually rehearsing a face, even after being prompted by a verbal description cue, may indeed improve recognition (Sporer, 1988). However, a growing body of research suggests that contrary to this intuition, efforts to describe a previously seen face can actually *impair* subsequent memory performance, at least under some circumstances.

In the original documentation of this counterintuitive effect of verbal description on face recognition (termed *verbal overshadowing*), Schooler and Engstler-Schooler (1990) showed participants a videotape of a bank robbery. Some participants were instructed to describe the robber in as much detail as possible while others engaged in an unrelated filler activity. Finally, all participants were shown a lineup containing the robber and seven foils. The results revealed that participants who had described the robber were markedly less accurate in recognizing him compared with no-description controls. Follow-up experiments by Schooler and Engstler-Schooler were largely consistent with the verbal overshadowing hypothesis that the negative effects of verbalization were due to a mismatch between the visual information or processes associated with the original experience and the verbal information or processes associated with the act of verbal description. For example, the negative effects of verbal description generalized to another type of nonverbal stimuli (i.e., colors), but not to more readily verbalized stimuli (i.e., the contents of what the robber said). Similarly, whereas verbal rehearsal repeatedly disrupted performance, visualizing the robber's face had no effect on subsequent identification.

Since its original demonstration, the verbal overshadowing phenomenon has been replicated numerous times (Dodson, Johnson, & Schooler, 1997; Fallshore & Schooler, 1995; Ryan & Schooler, 1998; Schooler, Ryan, & Reder, 1996; Sporer, 1989). At the same time, however, it has also failed to replicate on a number of occasions (Lovett, Small, & Engstrom, 1992; Yu & Geiselman, 1993). A meta-analysis of the verbal overshadowing effect was recently conducted by Meissner and Brigham (2001a). Across a sample of 15 studies (29 effect size comparisons; $N = 2018$), Meissner and Brigham observed a small, yet significant, verbal overshadowing effect ($Zr = -.12$) demonstrating that participants who described a target face were 1.27 times more likely to later *misidentify* the face from a lineup recognition task when compared with participants who did not generate a description prior to identification.

Although the verbal overshadowing effect is a reliable phenomenon, it nevertheless appears to be somewhat fragile. Moreover, while research following the original demonstration of verbal overshadowing is largely (if not entirely) consistent with the claim that it is associated with discrepancies between the modality of the original visual encoding, the precise mechanism responsible for the effect remains an issue of some con-

tention. We briefly review the research surrounding this topic and then consider the merits of several current explanations. As will be seen, there is compelling evidence in support of each of the primary accounts, yet no single explanation can accommodate all of the extant findings suggesting that multiple mechanisms may be involved.

Recoding Interference. In their original account of the verbal overshadowing effect, Schooler and Engstler-Schooler (1990) proposed that it results from recoding interference in which "the verbalization of a visual memory can foster the formation of a nonveridical verbally biased representation corresponding to the original stimulus" (p. 62). Such an account generally explains why the overshadowing effect is exclusively observed with nonverbal stimuli such as faces that are difficult to put into words, but not with stimuli that are more easily described.

More recently, Meissner and his colleagues (Meissner, 2002; Meissner et al., 2001) have provided additional support for the recoding interference account by demonstrating that the influence of verbalization is mediated by the amount of incorrect descriptors that participants are encouraged to generate. Specifically they found that verbal disruption was maximized when participants were "forced" to provide elaborate descriptions of the face. Under such forced recall conditions, Meissner and colleagues (2001) found that participants were more likely to include erroneous elements in their descriptions and subsequently demonstrated verbal overshadowing in their poor performance on a lineup identification task (27% accuracy) when compared with participants in a no-description control condition (52% accuracy). In contrast, another group of participants were warned to provide very accurate descriptions and not to guess at any particular features. Those in this warning condition actually demonstrated verbal enhancement (63% accuracy) when compared with participants in the control or forced conditions. Meissner and colleagues have replicated this "instructional bias" effect in several studies (Meissner, 2002; Meissner et al., 2001; see also Finger & Pezdek, 1999; MacLin, Tapscott, & Malpass, 2002) and have found that the effect persists despite delays of 30 minutes or 1 week, despite instructions to source monitor, and across repeated attempts at recall prior to identification. Taken together, these results suggest that extensive verbalization can lead to the production of a self-generated misinformation effect whereby participants are misled by the erroneous details present in their own descriptions. Further support of this account has also come from a moderator analysis conducted by Meissner and Brigham (2001a) demonstrating that variations in the reliability of verbal overshadowing studies could be reconciled by differences in the procedure used by various researchers. In particular, studies that utilized elaborative description procedures led to more reliable verbal overshadowing effects than those that utilized a standard free-recall procedure.

Transfer Inappropriate Processing Shift. Although the recoding interference account nicely accommodates many verbal overshadowing findings, there are some results that it does not easily handle (for a review see Schooler, Fiore, & Brandimonte, 1997; Schooler, 2002). First, whereas a relationship between verbalization quality and recognition performance has been observed in some studies (e.g., Finger & Pezdek, 1999,

Meissner, 2002; Meissner et al., 2001), other studies have failed to find such a relationship (e.g., Schooler & Engstler-Schooler, 1990). If verbal overshadowing is due to inaccuracies present in the verbal description, then, as Meissner and others have noted, such inaccuracies should be predictive of performance. The failure to find such a relationship across all studies, regardless of the type of recall instructions or task provided, is therefore potentially problematic for this account (Schooler, 2002).

A second problem for the recoding interference account involves studies demonstrating that verbalization can interfere with the recognition of other nonverbalized faces. For example, Dodson and colleagues (1997) presented participants with two faces (a male and a female face) and then had them describe just one of them. On a subsequent recognition test, they observed that verbalization interfered with the recognition of the nonverbalized face as much as it did with the verbalized face. Additional studies have demonstrated that even describing a parent's face from memory can interfere with recognition of a recently encoded (and unrelated) face. More recently, Brown and Lloyd-Jones (2002, 2003) have introduced a novel overshadowing paradigm in which participants are asked to encode a series of faces. Half of the participants are then asked to provide a description of the final face they viewed, and the second group of participants is asked to complete an unrelated filler task. Thereafter, all participants are provided with a recognition test in which a series of faces are shown to them (both novel faces and those from the study set). Brown and Lloyd-Jones have consistently found that describing the final face produces a verbal overshadowing effect in the recognition of all faces from the study set.

If verbal overshadowing is the product of relying on an inaccurate verbal code, then it is hard to understand why verbalization would have comparable effects when the face in question was itself never actually verbalized. Given these concerns, Schooler and his colleagues have suggested an alternative to the recoding interference account, originally termed "transfer inappropriate retrieval" (Schooler et al., 1997) but subsequently renamed "transfer inappropriate processing shift" (TIPS) (Schooler, 2002), based upon evidence that *retrieval* per se may not be a critical component of the process. According to Schooler and colleagues, verbal descriptions may induce a general processing shift that dampens the subsequent application of nonverbal configural processes. In effect, verbal description causes participants to become "stuck" in a verbal mode of processing faces, which is then applied (inappropriately transferred) to the recognition test, resulting in disruption.

The TIPS account nicely accommodates the basic finding that verbalization impairs recognition of nonverbal stimuli (such as faces), but not stimuli that are easily verbalizable (as only the former would be disrupted by an excessive focus on verbal processing). It also accounts for the findings that verbalizing one face can interfere with recognition of a different face (because of the general nature of the processing shift). Finally, TIPS is consistent with the influence of other manipulations (e.g., focusing on individual elements of composite figures) that disrupt face recognition performance (Macrae & Lewis, 2002) and provides a useful way of conceptualizing a variety of situations in which the engagement in one task can impair performance on subsequent tasks. At the same time, however, it does not offer a simple account of why a relation-

ship is sometimes observed between the quality of verbal descriptions and recognition performance.

Criterion Shifts. Until recently, the debate regarding the mechanisms underlying the negative effects of verbal description on face recognition were limited to the recoding interference and transfer inappropriate processing accounts. However, a third account has been suggested in which verbalization is said to more simply induce a criterion shift such that individuals who provide a description are subsequently less likely to make a positive identification (irrespective of accuracy). In a target-present lineup (used by the majority of researchers investigating the verbal overshadowing effect), such a shift would lead to a greater frequency of misses and thus to reduced accuracy. In testing this hypothesis, Clare and Lewandowsky (2004) found that verbal description of a previously presented face impaired performance on suspect present lineups when participants were provided a "not present" option, but not when they were forced to select from among the faces presented. Moreover, on a target-absent lineup, verbalization actually improved performance (being more cautious necessarily leads to less false identifications)—a finding that the authors note is not predicted by either the recoding interference or TIPS accounts.

While representing an important additional account of verbal overshadowing effects, Clare and Lewandowsky (2004) acknowledge that this approach cannot explain all the extant findings. Specifically, a number of studies have found verbal overshadowing effects with paradigms that either did not include a "not present" option (e.g., Fallshore & Schooler, 1995) or assessed performance on target-absent lineups (e.g., Meissner, 2002). In addition, the recognition paradigm introduced by Brown and Lloyd-Jones (2002, 2003) permitted the calculation of signal detection measures of discrimination and response criterion, but found an overshadowing effect on the former measure. Taken together, these findings prove difficult for a criterion shift account and encourage further research on the precise mechanism of the verbal overshadowing effect.

Summary of Verbal Overshadowing Findings. In the end it seems that all three current accounts of the negative effects of verbal description on face recognition have merit. Under some conditions, such as when individuals provide elaborate descriptions of a face and a relationship between description quality and recognition accuracy exists, it seems quite likely that verbalization produces a self-generated misinformation effect in which participants rely upon their erroneous description at the expense of their more veridical visual memory. Under other conditions, particularly when no relationship between description performance and recognition accuracy is observed and/or when verbalization is observed to impair the recognition of faces other than those described, it seems likely that verbalization induces a transfer inappropriate processing shift, whereby featural processing operations are inappropriately applied to a recognition test that would be better served by nonverbal, configural processes. Under still other situations, particularly when not present options are included, and the negative effects of verbalization are limited to increased misses, a criterion shift may be in operation. Clearly future research is needed to sort out more precisely when each of these respec-

tive mechanisms may be at play. Nevertheless, such research seems greatly warranted, given that verbal description is an inherent element in many eyewitness situations, and that understanding the precise mechanisms by which such descriptions can impair memory is certain to be critical to minimizing the negative effects that such descriptions might otherwise have. In the meantime, investigators should be cautioned against encouraging elaborate descriptions of a perpetrator, so as to minimize the effects of self-generated misinformation on later identification.

CONCLUSIONS AND FUTURE DIRECTIONS

A pervasive theme of research on eyewitness performance is that memory is not particularly reliable. Unfortunately, this theme appears to be particularly pronounced in the context of person descriptions. Person descriptions tend to be vague and nondiscriminative and are susceptible to many of the sources of error that plague other forms of eyewitness memory (e.g., the effects of arousal, poor encoding conditions, misinformation, declines with age, etc.). At the same time, there appear to be some aspects of person descriptions that are uniquely problematic. For example, whereas in general it is useful for witnesses to generate as much information about a witnessed event as possible (e.g., Fisher et al., 1989), in the context of person description, encouraging people to spend extensive time generating their descriptions can actually impair face recognition (Finger & Pezdek, 1999) and result in the generation of a greater proportion of inaccurate details (Meissner et al., 2001).

Although much has been learned about person description, there is still more that needs to be discerned. Theoretically, an important area for future research is to further flesh out the shared and unique processes that contribute to individuals' ability to recognize as opposed to describe faces. A variety of converging lines of evidence suggest that person descriptions may draw on processes that are distinct from those involved in face recognition. Whereas face recognition benefits from focusing on the global qualities of a face (Farah et al., 1998), face description benefits more from consideration of individual features (Wells & Turtle, 1987). Similarly, whereas face recognition consistently reveals an own-race advantage (a process known to rely on configural processing), face description has generally failed to show such a difference (Meissner & Brigham, 2001b; Sporer, 2001a, 2001b). These findings, in conjunction with a rather low or inconsistent relationship between the quality of face descriptions and recognition performance, as well as the verbal overshadowing phenomenon, suggest that face recognition and face description may rely on fundamentally different processes.

From this perspective it appears that future research might benefit from more precisely delineating the distinct processes contributing to person description versus recognition and explicating the behavioral and neurocognitive underpinnings of those processes. For example, recent research has found that face recognition performance is impaired if, between encoding and test, participants are shown large letters composed of small letters and are asked to attend to the smaller letters—a procedure believed to promote featural processing (Macrae & Lewis, 2002). However, what would be the effect of

such a manipulation on person description? Given the hypothesis that person description relies more on featural processing, it seems quite plausible that although a focus on local processing impairs face recognition, it may actually improve face description! It has also been observed that focusing on large letters in this task can enhance face recognition; however, according to the current perspective, a configural process might actually impair person description. It also seems quite plausible that person description and face recognition may differentially draw upon separate areas of the brain, with face recognition relying more on the nonverbal operations associated with the right hemisphere (Leehey, Carey, Diamond, & Cahn, 1978) and face description relying more on the verbal operations associated with the left hemisphere (Hellige, 1993). Further investigation of the unique and sometimes conflicting processes associated with person recognition and description may be crucial to enhancing our theoretical understanding of these two critical elements of eyewitness memory.

In addition to suggesting important theoretical directions for future research, the present analysis also points to some critical applied issues that must be resolved if we are to maximize the efficacy of person descriptions in eyewitness contexts. As noted, it appears that the value of person descriptions critically depends upon how much information individuals are required to generate, with extensive descriptions leading to both more inaccurate and more disruptive descriptions. However, determining the precise amount of information that will lead to maximum description quality has yet to be determined. Exactly how much information should witnesses be asked to provide? If details are not spontaneously offered, should they be probed for? And if so, which details are acceptable to inquire about, and which details may lead to elaborative interference? If a witness does offer an extensive description that is potentially more riddled with inaccuracies, are there some details (e.g., hair color) that might be more likely to be accurate than others (e.g., shape of face)? Are the details that are generated first more likely to be accurate than those generated later, and, if so, can the utility of person descriptions be enhanced by differentially emphasizing details that are more likely to be accurate from those that are more suspect? Although clearly there is much more that we need to research, we are in a far better position to know when and how to use this critical source of eyewitness information by recognizing the unique issues that affect person description processes.

ACKNOWLEDGMENTS

The writing of this manuscript was supported by a grant from the National Science Foundation to the first author (CAM) and a grant from the *Deutsche Forschungsgemeinschaft* (German Science Foundation) to the second author (SLS).

REFERENCES

Bahrick, H. (2000). Long-term maintenance of knowledge. In E. Tulving & F. I. M. Craik, *The Oxford handbook of memory* (pp. 347–362). New York: Oxford University Press.

Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (Vol. 2, pp. 396–401). London: Wiley.

Bornstein, B. H., Liebel, L. M., & Scarberry, N. C. (1998). Repeated testing in eyewitness memory: A means to improve recall of a negative emotional event. *Applied Cognitive Psychology, 12*, 119–131.

Brown, W. (1923). To what extent is memory measured by a single recall trial. *Journal of Experimental Psychology, 6*, 377–382.

Brown, C., & Lloyd-Jones, T. J. (2002). Verbal overshadowing in a multiple face presentation paradigm: Effects of description instruction. *Applied Cognitive Psychology, 16*, 873–885.

Brown, C., & Lloyd-Jones, T. J. (2003). Verbal overshadowing of multiple face and car recognition: Effects of within- versus across-category verbal descriptions. *Applied Cognitive Psychology, 17*, 183–201.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition, 20*, 277–290.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.

Chance, J. E., & Goldstein, A. G. (1996). The other-race effect and eyewitness identification. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 153–176). Mahwah, NJ: Lawrence Erlbaum Associates.

Christianson, S. (1984). The relationship between induced emotional arousal and amnesia. *Scandinavian Journal of Psychology, 25*, 147–160.

Christianson, S. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112*, 284–309.

Clare, J., & Lewandowsky, S. (2004). Verbalizing facial memory: Criterion effects in verbal overshadowing. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 739–755.

Clifford, B. R., & Hollin, C. R. (1981). Effects of the type of incident and the number of perpetrators on eyewitness memory. *Journal of Applied Psychology, 66*, 364–370.

Cohen, G., & Faulkner, D. (1989). Age differences in source forgetting: Effects on reality monitoring and on eyewitness testimony. *Psychology & Aging, 4*, 10–17.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). Improving the reliability of eyewitness identifications: Putting context into context. *Journal of Applied Psychology, 72*, 629–637.

Cutshall, J., & Yuille, J. C. (1989). Field studies of eyewitness memory of actual crimes. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 97–124). New York: Springer.

Darley, C. F., & Glass, A. L. (1971). Effects of rehearsal and serial list position on recall. *Journal of Experimental Psychology, 93*, 83–89.

Davies, G. M. (1996). Children's identification evidence. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 233–258). Mahwah, NJ: Lawrence Erlbaum Associates.

Davies, G., Tarrant, A., & Flin, R. (1989). Close encounters of the witness kind: Children's memory for a simulated health inspection. *British Journal of Psychology, 80*, 415–429.

Deffenbacher, K. A. (1983). The influence of arousal on reliability of testimony. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 235–251). Chichester, England: Wiley.

Deffenbacher, K. A. (1994). Effects of arousal on everyday memory. *Human Performance, 7*, 141–161.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law & Human Behavior, 28*, 687–706.

Dent, H. (1982). The effects of interviewing strategies on the results of interviews with child witnesses. In A. Trankell (Ed.), *Reconstructing the past* (pp. 279–298). Stockholm: Norstedt.

Dent, H., & Stephenson, G. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social & Clinical Psychology, 18*, 41–51.
Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition, 25*, 129–139.
Dunning, D., & Stern, L. B. (1992). Examining the generality of eyewitness hypermnesia: A close look at time delay and question type. *Applied Cognitive Psychology, 6*, 643–657.
Ebbesen, E. B., & Rienick, C. B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology, 5*, 745–762.
Ellis, H. D. (1992). The development of face processing skills. *Proceedings of the Royal Society, Series B, 335*, 105–111.
Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of white and black faces by white and black subjects. *International Journal of Psychology, 10*, 119–123.
Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1980). The deterioration of verbal descriptions of faces over different delay intervals. *Journal of Police Science & Administration, 8*, 101–106.
Fallshore, M., & Schooler, J. W. (1995). The verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 1608–1623.
Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498.
Finger, K., & Pezdek, K. (1999). The effect of verbal description on face identification accuracy: "Release from verbal overshadowing." *Journal of Applied Psychology, 84*, 340–348.
Fisher, G. H., & Cox, R. L. (1975). Recognizing human faces. *Applied Ergonomics, 6*, 104–109.
Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing.* Springfield, IL: Charles C. Thomas.
Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of a crime. *Journal of Applied Psychology, 74*, 722–727.
Flin, R. H., & Shepherd, J. W. (1986). Tall stories: Eyewitnesses' ability to estimate height and weight characteristics. *Human Learning, 5*, 29–38.
Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17*, 533–543.
Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L. A., Sullivan, S. J., Avetissian, I. V., et al. (1984). Enhancement of eyewitness memory: An empirical evaluation of a cognitive interview. *Journal of Police Science and Administration, 12*, 74–80.
Glenberg, A., & Adams, F. (1978). Type I rehearsal and recognition. *Journal of Verbal Learning and Verbal Behavior, 17,* 455–163.
Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law & Human Behavior, 10*, 317–332.
Grass, E., & Sporer, S. L. (1991, March). *Richtig oder falsch? Zur Vorhersage von Identifizierungsleistungen durch weitere Aussagen von Zeugen* [Correct or false? Post-dicting eyewitness identification accuracy from verbal statements]. Paper presented at the 33rd Tagung experimentell arbeitender Psychologen in Giesen, Germany.
Hellige, J. B. (1993). Unity of thought and action: Varieties of interaction between the left and right cerebral hemispheres. *Current Directions in Psychological Science, 2*, 21–25.
Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General,* 127, 251–268.
Janssen, J. P., & Horowski, A. C. (1980). Schueler schaetzen Koerpergroessen: Akzentuierungstendenz als kognitiver Stil der Personenwahrnehmung [Children estimate height: Accentuation tendencies as a cognitive style in person perception]. *Zeitschrift für Entwicklungspsychologie und Paedagogische Psychologie, 10*, 167–176.
Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology, 65*, 190–193.

Kleinsmith, L. J., & Kaplan, S. (1964). Interaction of arousal and recall interval in nonsense syllable paired-associate learning. *Journal of Experimental Psychology, 67*, 124–126.

Koehnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime & Law, 5*, 3–27.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490–517.

Kuehn, L. L. (1974). Looking down a gun barrel: Person perception and violent crime. *Perceptual & Motor Skills, 39*, 1159–1164.

Laughery, K. R., Duval, C., & Wogalter, M. S. (1986). Dynamics of facial recall. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 373–387). Dordrecht: Martinus Nijhoff.

Leehey, S., Carey, S., Diamond, R., & Cahn, A. (1978). Upright and inverted faces: The right hemisphere knows the difference. *Cortex, 14*, 411–419.

Leippe, M. L., Romanczyk, A., & Manion, A. P. (1991). Eyewitness memory for a touching experience: Accuracy differences between child and adult witnesses. *Journal of Applied Psychology, 76*, 367–379.

Lindholm, T. (2005). Own-age biases in verbal person memory. *Memory, 13*, 21–30.

Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law & Human Behavior, 18*, 527–541.

Lipton, J. P. (1977). On the psychology of eyewitness testimony. *Journal of Applied Psychology, 62*, 90–95.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560–572.

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E., & Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition, 10*, 318–323.

Loftus, E. F., & Greene, E. (1980). Warning: Even memory for faces may be contagious. *Law & Human Behavior, 4*, 323–334.

Loftus, E. F., & Ketcham, K. E. (1983). The malleability of eyewitness accounts. In S. M. A. Lloyd-Bostock and B. R. Clifford (Eds.) *Evaluating witness evidence* (pp. 159–171). New York: John Wiley & Sons.

Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society, 5*, 86–88.

Loftus, G. R. (1985). Picture perception: Effects luminance on available information and information extraction rate. *Journal of Experimental Psychology: General, 114*, 342–356.

Lovett, S. B., Small, M. Y., & Engstrom, S. A. (1992, November). *The verbal overshadowing effect: Now you see it, now you don't*. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, MO.

MacLeod, M. D., & Shepherd, J. W. (1986). Sex differences in eyewitness reports of criminal assaults. *Medicine, Science & the Law, 26*, 311–318.

MacLin, O. H., Tapscott, R. L., & Malpass, R. S. (2002). The development of a computer system to collect descriptions of culprits. *Applied Cognitive Psychology, 16*, 937–945.

Macrae, C. N., & Lewis, H. L. (2002). Do I know you?: Processing orientation and face recognition. *Psychological Science, 13*, 194–196.

Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law & Human Behavior, 3*, 295–305.

Meissner, C. A. (2002). Applied aspects of the instructional bias effect in verbal overshadowing. *Applied Cognitive Psychology, 16*, 911–928.

Meissner, C. A., & Brigham, J. C. (2001a). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology, 15*, 603–616.

Meissner, C. A., & Brigham, J. C. (2001b). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law, 7*, 3–35.

Meissner, C. A., Brigham, J. C., & Kelley, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition*, 29, 176–186.

Muensterberg, H. (1908). *On the witness stand. Essays on psychology and crime*. New York: Doubleday, Page.

Murray, J. B. (1986). Marijuana's effects on human cognitive functions, psychomotor functions, and personality. *Journal of General Psychology, 113,* 23–55.

Neil v. Biggers, 409 U.S. 188 (1972).

Payne, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin, 101*, 5–27.

Pickel, K. L. (1998). Unusualness and threat as possible causes of "weapon focus." *Memory, 6*, 277–295.

Pickel, K. L. (1999). The influence of context on the "weapon focus" effect. *Law & Human Behavior, 23*, 299–311.

Pigott, M., & Brigham, J. C. (1985). Relationship between accuracy of prior description and facial recognition. *Journal of Applied Psychology, 70,* 547–555.

Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identification of strangers by youth and adult eyewitnesses. *Journal of Applied Psychology, 88*, 315–323.

Pryke, S., Lindsay, R. C. L., & Pozzulo, J. D. (2000). Sorting mug shots: Methodological issues. *Applied Cognitive Psychology, 14*, 81–96.

Read, J. D., Hammersley, R., Cross-Calvert, S., & McFadzen, E. (1989). Rehearsal of faces and details in action events. *Applied Cognitive Psychology, 3*, 295–311.

Read, J. D., Yuille, J. C., & Tollestrup, P. (1992). Recollections of a robbery: Effects of arousal and alcohol upon recall and person identification. *Law & Human Behavior, 16*, 425–446.

Reinhardt-Rutland, A. H. (1986). Note on nonveridical visual perception and pedestrian accidents at night. *Perceptual & Motor Skills, 63*, 371–374.

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology, 80*, 313–331.

Roediger, H. L., & Challis, B. H. (1989). Hypermnesia: Improvements in recall with repeated testing. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 175–199). Hillsdale, NJ: Lawrence Erlbaum Associates.

Roediger, H. L., Jacoby, J. D., & McDermott, K. B. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory and Language, 35,* 300–318.

Roediger, H. L., McDermott, K. B., & Goff, L. M. (1997). Recovery of true and false memories: Paradoxical effects of repeated testing. In M. A. Conway (Ed.), *Recovered memories and false memories* (pp. 118–149). Oxford: Oxford University Press.

Roediger, H. L., Wheeler, M. A., & Rajaram, S. (1993). Remembering, knowing, and reconstructing the past. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 97–134). San Diego: Academic Press.

Ryan, R. S., & Schooler, J. W. (1998). Whom do words hurt?: Individual differences in susceptibility to verbal overshadowing. *Applied Cognitive Psychology, 12*, 105–126.

Sayette, M. A. (1999). Cognitive theory and research. In K. Leonard & H. Blume (Eds.), *Psychological theories of drinking and alcoholism* (2nd ed., pp. 247–291). New York: Guilford press.

Saywitz, K. J. (1995). Improving children's testimony: The question, the answer, and the environment. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 113–140). London: Sage.

Schooler, J. W. (1998). The distinctions of false and fuzzy memories. *Journal of Experimental Child Psychology, 71*, 130–143.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology, 16*, 989–997.

Schooler, J. W., & Eich, E. (2000). Memory for emotional events. In E. Tulving & F. I. M. Craik, *The Oxford handbook of memory* (pp. 379–392). New York: Oxford University Press.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22*, 36–71.

Schooler, J. W., Fiore, S. M., & Brandimonte, M. A. (1997). At a loss *from* words: Verbal overshadowing of perceptual memories. In D. Medin's (Ed.), *Handbook of learning and motivation* (Vol. 37). Orlando, FL: Academic Press.

Schooler, J. W., Foster, R. A., & Loftus, E. F. (1988). Some deleterious consequences of the act of recollection. *Memory and Cognition, 16*, 243–251.

Schooler, J. W., & Loftus E. F. (1993). Multiple mechanisms mediate individual differences in eyewitness accuracy and suggestibility. In H. W. Reese and J. M. Puckett (Eds.), *Mechanisms of practical cognition* (pp. 177–203). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schooler, J. W., Ryan, R. S., & Reder, L. M. (1996). The costs and benefits of verbalization. In D. Herrmann, M. Johnson, C. McEvoy, C. Hertzog, & P. Hertels (Eds.), *Basic and applied memory: New findings* (pp. 51–65). Mahwah, NJ: Lawrence Erlbaum Associates.

Searcy, J. H., Bartlett, J. C., Memon, A., & Swanson, K. (2001). Aging and lineup performance at long retention intervals: Effects of metamemory and context reinstatement. *Journal of Applied Psychology, 86*, 207–214.

Shaw, J. S., Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin and Review, 2*(2), 249–253.

Shaw, J. S., Garven, S., & Wood, J. M. (1997). Co-witness information can have immediate effects on eyewitness memory reports. *Law & Human Behavior, 21*, 503–523.

Shepherd, J. W. (1983). Identification after long delays. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 173–187). Chichester: Wiley.

Shepherd, J. W. (1986). An interactive computer system for retrieving faces. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 398–409). Dordrecht/Boston/Lancaster: Martinus Nijhoff.

Shepherd, J. W., & Deregowski, J. B. (1981). Races and faces: A comparison of the responses of Africans and Europeans to faces of the same and different races. *British Journal of Social Psychology, 20*, 125–133.

Shepherd, J. W., & Ellis, H. D. (1996). Face recall—Methods and problems. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 87–115). Mahwah, NJ: Lawrence Erlbaum Associates.

Shepherd, J. W., Ellis, H. D., & Davies, G. M. (1977). *Perceiving and remembering faces*. Technical report to the Home Office under Contract no. POL/73/1675/24/1.

Shepherd, J. W., Ellis, H. D., & Davies, G. M. (1982). *Identification evidence: A psychological examination*. Aberdeen: Aberdeen University Press.

Sporer, S. L. (in press). Person descriptions as retrieval cues: Do they really help? *Psychology, Crime, and Law.*

Sporer, S. L. (1988). Long-term improvement of facial recognition through visual rehearsal. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 182–188). London: Wiley.

Sporer, S. L. (1989). Verbal and visual processes in person identification. In H. Wegener, F. Lösel, & J. Haisch (Eds.), *Criminal behavior and the criminal justice system* (pp. 303–324). New York/Berlin: Springer.

Sporer, S.L. (1991). Encoding strategies and the recognition of human faces. *Journal of Experimental Psychology: Human Learning, Memory and Cognition, 17,* 323–333.

Sporer, S. L. (1992a, March). *An archival analysis of person descriptions.* Paper presented at the Biennial Meeting of the American Psychology-Law Society in San Diego, California.

Sporer, S. L. (1992b). Post-dicting eyewitness accuracy: Confidence, decision-times and person descriptions of choosers and non-choosers. *European Journal of Social Psychology, 22*, 157–180.
Sporer, S. L. (1996a). Describing others: Psychological issues. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 53–86). Mahwah, NJ: Lawrence Erlbaum Associates.
Sporer, S. L. (1996b). Experimentally induced person mix-ups through media exposure and ways to avoid them. In G. M. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology and law: Advances in research* (pp. 64–73). Berlin: De Gruyter.
Sporer, S. L. (2001a). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, & Law, 7*, 36–97.
Sporer, S. L. (2001b). The cross-race bias: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and Law, 7,* 170–200.
Sporer, S. L., & Cutler, B. L. (2003). Identification evidence in Germany: Common sense assumptions, empirical evidence, guidelines, and judicial practices. In P. J. van Koppen & S. D. Penrod (Eds.), *Adversarial vs. inquisitorial justice: Psychological perspectives on criminal justice systems* (pp. 191–208). New York: Plenum.
Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law & Human Behavior, 16*, 413–424.
Stern, L. W. (1902). Zur Psychologie der Aussage [Psychology of report]. *Zeitschrift für die gesamte Strafrechtswissenschaft, 22*, 315–370.
Turtle, J. W., & Yuille, J. C. (1994). Lost but not forgotten details: Repeated eyewitness recall leads to reminiscence but not hypermnesia. *Journal of Applied Psychology, 79*, 260–271.
van Koppen, P., & Lochun, S. (1997). Portraying perpetrators: the validity of offender descriptions by witnesses. *Law and Human Behavior, 21*, 661–685.
Wagstaff, G. F., MacVeigh, J., Boston, R., Scott, L., Brunas-Wagstaff, J., & Cole, J. (2003). Can laboratory findings on eyewitness testimony be generalized to the real world? An archival analysis of the influence of violence, weapon presence, and age on eyewitness accuracy. *Journal of Psychology: Interdisciplinary & Applied, 137*, 17–28.
Warnick, D. H., & Sanders, G. S. (1980). The effects of group discussion on eyewitness accuracy. *Journal of Applied Social Psychology, 10,* 249–259.
Wells, G. L. (1985). Verbal descriptions of faces from memory: Are they diagnostic of identification accuracy. *Journal of Applied Psychology, 70*, 619–626.
Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology, 66*, 682–687.
Wells, G. L., & Turtle, J. W. (1987). What is the best way to encode faces? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research & issues* (pp. 163–168). New York: Wiley & Sons.
Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25*, 731–739.
Wogalter, M. S. (1991). Effects of post-exposure description and imaging on subsequent face recognition performance. *Proceedings of the Human Factors Society, 35*, 575–579.
Wogalter, M. S. (1996). Describing faces from memory: Accuracy and effects on subsequent recognition performance. *Proceedings of the Human Factors and Ergonomics Society, 40*, 536–540.
Yarmey, A. D. (1986). Verbal, visual, and voice identification of a rape suspect under different levels of illumination. *Journal of Applied Psychology, 71*, 363–370.
Yarmey, A. D. (1993). Adult age and gender differences in eyewitness recall in field settings. *Journal of Applied Social Psychology, 23*, 1921–1932.
Yarmey, A. D. (1996). Age and eyewitness memory. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 259–278). Mahwah, NJ: Lawrence Erlbaum Associates.

Yarmey, A. D. (2004). Eyewitness recall and photo identification: A field experiment. *Psychology, Crime & Law, 10*, 53–68.

Yarmey, A. D., Jacob, J., & Porter, A. (2002). Person recall in field settings. *Journal of Applied Social Psychology, 32*, 2354–2367.

Yarmey, A. D., & Kent, J. (1980). Eyewitness identification by elderly and young adults. *Law & Human Behavior, 4*, 359–371.

Yarmey, A. D., & Morris, S. (1998). The effects of discussion on eyewitness memory. *Journal of Applied Social Psychology, 28*, 1637–1648.

Yarmey, A. D., & Yarmey, M. (1997). Eyewitness recall and duration estimates in field settings. *Journal of Applied Social Psychology, 27*, 330–344.

Yarmey, A. D., Yarmey, A. L., & Yarmey, M. J. (1994). Face and voice identifications in showups and lineups. *Applied Cognitive Psychology, 8*, 453–464.

Yu, C. J., & Geiselman, R. E. (1993). Effects of constructing identi-kit composites on photospread identification performance. *Criminal Justice & Behavior, 20*, 280–292.

Yuille, J. C. (1986). Meaningful research in the police context. In J. C. Yuille (Ed.), *Police selection and training: The role of psychology* (pp. 225–246). Dordrecht, the Netherlands: Martinus Nijhoff.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71*, 291–301.

Yuille, J. C., & Tollestrup, P. A. (1990). Some effects of alcohol on eyewitness memory. *Journal of Applied Psychology, 75*, 268–273.

Yuille, J. C., Tollestrup, P. A., Marxsen, D., Porter, S., & Herve, H. F. M. (1998). An exploration on the effects of marijuana on eyewitness memory. *International Journal of Law & Psychiatry, 21*, 117–128.

Zima, H., & Zeiner, W. (1982). Das Versuchsprojekt "Sigma." Neue Wege zur Erfassung und Auswertung von Signalementdaten [Project "Sigma": New way of gathering and evaluating person descriptions]. *Kriminalistik, 36*, 593–596.