Effects of stress induced by a simulated shooting on recall by police and citizen witnesses

CLAUDIA J. STANNY
University of West Florida

THOMAS C. JOHNSON
Mississippi State University

Two experiments address the accuracy of citizen and police witnesses in recalling stressful events. Participants saw 2 training scenarios. Experiment 1 evaluated the effects of active engagement and the presence of a simulated shooting on police officers' memory for details. Police officers recalled significantly fewer details from the scenario in which a shooting occurred. Active engagement did not influence recall. Experiment 2 compared the memory performance of police and citizen witnesses. Manipulation checks were added to assess the arousal produced by a shooting, and control procedures were added to evaluate the memorability of specific scenarios. Police and citizen witnesses did not differ reliably in accuracy. The perpetrator was remembered less well than the weapon when a simulated shooting occurred, and one scenario was significantly more difficult to remember than the other. The need to replicate eyewitness research with a variety of materials is discussed.

When eyewitness testimony is presented during a trial, jurors must evaluate the accuracy and credibility of the witness (Undeutsch, 1984). Eyewitness testimony can have a tremendous impact on a jury (Cutler, Penrod, & Dexter, 1989; Loftus, 1979). However, witnessed events often occur under less than optimal viewing conditions. Moreover, witnesses to crimes often experience high levels of stress or emotional arousal. Much research has been directed at understanding the effects of viewing conditions and emotional arousal on eyewitness accuracy (Buckhout, 1974; Christianson, 1992; Deffenbacher, 1991; Egeth, 1993, 1994; Heuer & Reisberg, 1992; Loftus, 1986; Narby, Cutler, & Penrod; 1996; Sporer, Malpass, & Koehnken, 1996; Wells, 1993; Yuille & Tollestrup, 1992). However, little research has been directed at comparing police and citizen witness accuracy. Yarmey (1986) reported that police, prosecutors, defense attorneys, university students, and citizens all believe that police are more perceptive of people and events in their environ-
ment than citizens. In addition, respondents believed that police are more accurate than citizens in identifying faces and expected that police would correctly recall more details about a suspect (Cutler & Penrod, 1995; Deffenbacher & Loftus, 1982). Is the superior credibility attributed to police witnesses justified?

Some theoretical accounts of memory predict that police, like citizens, might remember traumatic events poorly. For example, Yarmey (1988) described perceptual distortions experienced by police when a shooting occurs that might impair their memory. These include tunnel vision and distorted time perception. In addition, police may divert their attention away from processing peripheral details when they are busy responding to rapidly unfolding events. This may result in “tunnel memory” (Safer, Christianson, Autry, & Österlund, 1998), in which peripheral information is forgotten and participants remember pictures as being more closely focused when they depict traumatic scenes.

On the other hand, police training and experience may give police superior observation skills that enable them to remember witnessed events more accurately than citizens. For example, police knowledge about criminal behavior may make them more attentive to details in the environment that are relevant to detecting criminal activity. Lindholm, Christianson, and Karlsson (1997) proposed that police experience and training may enhance their performance in several ways. Police experience may provide domain-specific knowledge that facilitates encoding and storing new information. Police knowledge of variables that influence eyewitness accuracy may lead them to adopt strategies to counteract these influences. Police who witness a crime in progress are also likely to be active participants in these events. Engelkamp (1997) showed that people remember activities they perform better than activities they imagine performing. Finally, police might experience lower physiological or psychological arousal during stressful events than citizens for a variety of reasons. Police work might attract people who are less reactive to stressful events. Experience and training might enable police to respond to criminal episodes with lower levels of stress. Finally, police may learn to use processing strategies that allow them to manage their stress and still engage in adequate encoding. The evidence relevant to these claims is addressed here.

With regard to police knowledge, questionnaire studies suggest that police are not more knowledgeable than other respondents about the effects of variables on the accuracy of eyewitness reports (Winterdyk, 1988; Yarmey, 1986). Although researchers have found reliable individual differences in skill at face recognition, people with greater skill do not come from any clearly identifiable professional group (Narby et al., 1996; Woodhead & Baddeley, 1981) Moreover, attempts to train people
to recognize faces more accurately have met with disappointing results (Woodhead, Baddeley, & Simmonds, 1979). Comparisons of police and nonpolice observers suggest that police and citizens are similar in their observation skills, except that police are more likely to report observing criminal activity when none is present (Tickner & Poulton, 1975).

Training and experience might be expected to enable police officers to focus attention on crime-related behavior and encode more information than citizens. Clifford and Richards (1977) compared the accuracy of police and citizens in recalling a target person after the person engaged them in a social interaction that was either 15 s long (asking the correct time) or more than 30 s long (asking directions). Police and citizens did not differ in their memory for the target person after the 15-s interaction, but police were significantly more accurate than citizens after the 30-s interaction. Clifford and Richards proposed that the task of giving directions poses fewer demands on attentional resources for police, who give directions more often than do citizens. Thus, they can presumably devote more attention to encoding details about the person.

Yuille (1984) compared police with citizen witnesses and found that police recalled more descriptive facts than did citizens. Recent work by Lindholm et al. (1997) suggests that police recall more information than citizens about the perpetrator and the weapon used in a crime. However, although police recalled more information overall than did citizens, this difference was not statistically reliable for free-recall or multiple-choice measures. Police and citizens were also equally accurate in a lineup identification task. Christianson, Karlsson, and Persson (1998) compared the ability of police, students, and teachers to remember information from a violent film of a simulated crime. Although police recalled more information about the perpetrator than did some of the citizen witness groups, they did not differ significantly from comparison groups in their memory for all types of information. In contrast, several reviews of the eyewitness literature concluded that there is little evidence that police recall witnessed events any more accurately than citizens (Deffenbacher, 1991; Narby et al., 1996; Yarmey, 1986). Yarmey (1986) concluded that police are no more accurate in the visual identification of faces and are no more perceptive of events in their environment than are other witnesses. The small number of studies that compare the accuracy of eyewitness reports given by police and citizen witnesses and the variability of the results obtained make additional comparisons of police and citizen eyewitness reports desirable.

Criminal episodes often are surprising and stressful events. During police training, officers learn to anticipate dangerous situations, remain calm in the face of adversity, and apply strategies to defuse volatile sit-
uations (Florida Department of Law Enforcement, 1990). Although repeated exposure to dangerous, stressful situations might blunt the intensity of emotional responses to these events, few empirical data have been provided to evaluate this assumption. Christianson et al. (1998) and Lindholm et al. (1997) found that police report experiencing less intense emotional responses when viewing violent stimulus materials than do citizens. However, they did not include assessments of memory for nonviolent scenarios in their research.

In one of the few studies that evaluated the effects of a stress manipulation on police eyewitness memory, Yuille, Davies, Gibling, Marxsen, and Porter (1994) found that police trainees recalled fewer details from stressful incidents in a role-playing exercise than from nonstressful incidents. However, this measure does not take the accuracy of these details into account. For example, participants can produce more details during recall by producing additional but erroneous details. When only correct details were considered (as a percentage of the total number of details produced), Yuille et al. found that police trainees produced a larger percentage of correct details from the stressful condition than from the nonstressful condition.

A major purpose of the experiments reported here was to compare the effects of stress induced by a simulated shooting on the accuracy of eyewitness reports obtained from citizen and police witnesses. Because the effects of experimental manipulations on the quality of eyewitness memory can depend on the realism of the witnessed events (Christianson, 1992; Yuille & Tollestrup, 1992), the scenarios used as stimulus materials were realistic video simulations from the Firearms Training System (Firearms Training System, Inc., 1990) library. The Firearms Training System (FATS) is a computerized simulation used for police training. It consists of a video projection system that displays digitized scenarios from a videodisc player. The system is designed to create a realistic training experience for police officers. Video images for the projected scenarios are approximately life-size. Scenarios are presented in color and with sound. Illumination in the training room is kept low to maximize the brightness and detail of projected images. The videos were recorded from the perspective of an officer intervening at the scene of a potentially dangerous situation. The images are intended to simulate the experience of seeing events unfold firsthand.

The training room includes a partial brick wall to provide cover when the officer believes that his or her safety is threatened. Officers are instructed to issue orders, move about, seek cover, and draw and fire their weapons if they decide that these actions are necessary to deal with the events depicted. Rapid access to digitized images on the videodisc en-
able the system to branch to different endings during presentation of a scenario. During training, this feature is used to vary the appropriateness of a shoot response for a given scenario and is used to train officers to make decisions about the use of lethal force. Trainees are issued a field weapon and instructed to fire at targets in the simulation when their judgment indicates that use of lethal force is necessary. Although the firearms used during training are not loaded with ammunition, they produce a loud report when fired and emit a laser signal that is detected by the simulation equipment. At the end of a scenario, the software provides data on weapon proficiency, including data on reaction time, number of shots fired, location of hits and missed shots, and an evaluation of the officer's weapon use judgment.

The realism of these training scenarios makes them particularly attractive as controlled experimental materials for eyewitness research. The branching feature of the FATS system was used in Experiment 2 to create two versions of scenarios that differed only in the final few seconds of activity: Scenarios ended in either a peaceful resolution of the initial conflict or an escalation to the use of lethal force by the participating officer. The content of the scenarios selected is described in detail the Methods section. Although the realism of these simulations might enhance ecological validity, it also makes assignment of citizen witnesses to active participation conditions undesirable. Thus, only police officers used the firearms during these studies and citizen witnesses (in Experiment 2) served only in bystander conditions. This constraint raises an important question for interpreting the results. Hosch and Bothwell (1990) compared the memory of victims and bystanders for a staged event. Although victims and bystanders were equally likely to select the perpetrator in a lineup, victims gave more accurate descriptions of the perpetrator than did bystanders. Similarly, Yuille et al. (1994) found that police trainees participating in role-playing incidents recalled more details than did observers. Christianson and Hübinnie (1993) also found that victims (tellers) remembered more details about a bank robbery than did bystanders (other employees and customers). However, these findings of differences in the accuracy of victims and bystanders might be confounded by differences in the point of view and the resulting amount of information available to victims and bystanders. The simulation videos used for the studies reported here, however, provided active participants and bystanders similar access to information in the witnessed event. Thus, Experiment 1 was conducted to compare the effects of active and passive participation on eyewitness accuracy when participants and bystanders have similar perspectives and when all witnesses have similar training and experience.
EXPERIMENT 1

METHOD

Participants

The participants in this study were 40 full-time, sworn law enforcement officers (4 women and 36 men) with the Pensacola, Florida, Police Department. The mean age of participants was 31 years, with a mean of 7.8 years of experience in law enforcement.

Materials

The scenarios were presented via the Firearms Training System (1990), a computerized, interactive, large-screen videodisc projection system designed for police training in weapon use and lethal force decisions. Scenarios were presented in color and with sound. The system was designed to maximize the realism of the trainee's perceptual experience. The training scenarios were filmed from the perspective of an officer intervening at the scene of a potentially dangerous situation. The camera shots were, in effect, the line of sight of the intervening officer. Video displays measured approximately 2.8 m wide and 2.1 m tall and were viewed from a distance of approximately 4 m. Participants used a weapon similar to their duty weapons. However, the weapon was fitted with a laser emitter that allowed the system to monitor when and where participants fired their weapons. The system was interactive and modified the video presentation so that perpetrators on the screen responded as if being shot when a laser "bullet" fired by the trainee intersected their screen images. For example, if an officer fired a lethal hit, the scenario ended. If the officer failed to shoot or if the shots were not lethal, the perpetrator could return fire. (However, in these experiments, all officers fired at appropriate times during the shoot condition, and this latter version of the scenario was never shown.)

Two training scenarios were selected from the FATS library for use as stimulus materials. The scenarios were similar in duration, complexity, and number of persons depicted. Both scenarios involved a man, a woman, and a police partner. The partner's voice was heard in both scenarios and the partner appeared briefly on the screen in one scenario. The action occurred next to a vehicle in both scenarios. A handgun appeared in both scenarios, although a knife also appeared in one scenario. Although both scenarios included perpetrators wielding guns, the perpetrator handled the gun in a nontargeting manner in the no-shoot scenario but aimed and fired the gun at the officer in the shoot scenario. Brief synopses of the two scenarios are given here.

**Domestic disturbance.** The domestic disturbance scenario depicted a woman standing next to a white Jeep station wagon, arguing with the driver, who could not be seen through the tinted windows. They were arguing over possession of the car. Off camera, a police partner asked what was going on. The woman complained that "he is trying to take my car." The car engine started, and the woman moved to the front of the car and pushed on the car. The police partner ordered the driver to stop the car, took the woman aside, and asked
the participant to watch the driver. As the woman and police partner exited, the driver opened the car door and exited from the car. He stood briefly near the front of the car, then reached into his back pocket. The driver pulled out a handgun with the barrel pointed downward, immediately put it on the ground, and resumed his stance. Although a weapon appeared in this scenario, the perpetrator handled it in a nonthreatening manner and officers were not expected to shoot during this sequence.

**Attempted abduction.** The attempted abduction scenario initially depicted only the interior of a parking structure, and a voiceover told the participant that the police had been summoned to investigate a report of a woman screaming. After a few seconds, a woman’s scream was heard. The camera view depicted a turn toward a part of the parking structure where a man and woman could be seen struggling next to a car. They were briefly out of view while the officers exited the car. The perpetrator was holding a knife to the woman’s throat. The partner ordered the perpetrator to release the woman. She was released and ran off camera. The police partner ordered the perpetrator to drop the knife, which he did. The perpetrator then pulled a gun from his pocket, aimed, fired once, and ran off camera (unless the participant fired a fatal shot, in which case the perpetrator fell to the ground). The explicitly threatening use of a gun by the perpetrator made a shoot response appropriate during this sequence.

Run time for the domestic disturbance scenario was 43 s. Run time for the attempted abduction scenario was somewhat longer (52 s), but this scenario also had a 14-s introductory sequence in which the patrol car drove inside the parking structure and the actors were not in view. Action segments were defined as the intervals during which active contact between police and actors took place. This interval began when voice commands were issued by the police partner and continued until the end of the scenario. In the domestic disturbance scenario this interval was 20 s; in the attempted abduction scenario this interval was 21 s. Within the action segments, a second action phase was defined, beginning when the victim exited and the action centered on the interaction between the perpetrator and the trainee/participant. This sequence lasted 11 s in the domestic disturbance scenario and 8 s in the attempted abduction scenario. The perpetrator was visible on screen for 11 s in the domestic disturbance scenario and 21 s in the attempted abduction scenario. The victim was visible on screen for 32 s in the domestic disturbance scenario and 10 s in the attempted abduction scenario. A gun was visible for 6 s in the domestic scenario. A knife was present for 6 s in the attempted abduction scenario and a gun was visible for 2 s.

A questionnaire was constructed to probe recall for details from the scenarios. Questions were worded so that they would apply to either scenario (e.g., “describe the victim’s shirt/blouse,” “describe the perpetrator’s pants/skirt/blouse,” “what was the make of the vehicle at which this incident occurred?”). The questionnaire contained nine questions about the victim, nine questions about the perpetrator, six questions about weapons, two questions about actions taken by the participant, and four questions about other details. Because the question about a second weapon was relevant only for one scenario, participants
were instructed to write N/A if they recalled that a question did not apply to a scenario and to leave the question blank if they recalled that the question was relevant but could not recall the detail. Thus, the question about a second weapon was scored as correct if the participant correctly indicated that this was not relevant to the domestic disturbance scenario. The questionnaire is provided in the Appendix.

Procedure

The participants were matched according to years of experience and age and participated in pairs. The participants were presented with a no-shoot scenario and a shoot scenario, with order of presentation counterbalanced across pairs. One participant from each pair was randomly assigned to the passive group, and the other participant was assigned to the active group. The passive participant was instructed to watch the scenarios and was placed in a position where he or she could watch both the scenarios and the behavior of the active participant. No active involvement with the scenarios was required.

The active participant was told to respond to the scenarios in a manner consistent with his or her training and experience. The participant was instructed to move, issue verbal commands, draw his or her weapon, and fire when appropriate during the scenarios. The shoot scenario (attempted abduction) was a training scenario in which active participants showing good police judgment should shoot their weapons. Shooting was not an appropriate response during the no-shoot scenario (domestic disturbance).

The participants believed they were engaged in an evaluation of the training system. They were not told in advance that their memory of the scenarios would be tested. However, after both scenarios were presented, the participants were asked to complete the questionnaires for recall of details from the two scenarios. They were later told that the study was actually an assessment of the effects of the scenarios on memory.

RESULTS

The level of violence-related arousal was manipulated by selecting scenarios that varied in terms of whether they warranted a decision to shoot. Consistent with this criterion, all 20 active participants fired their training revolvers during the shoot scenario and only 2 participants fired during the no-shoot scenario.

Details supplied in response to each item in the questionnaire were scored as either correct or incorrect. Omissions were scored as incorrect unless the participant correctly indicated that the question was not relevant to a scenario. Proportions of correct recall were computed for each type of detail, based on the number of questions of a given type answered correctly.

Proportions of correct recall scores were analyzed with a $2 \times 2 \times 5$ mixed ANOVA in which mode of involvement (active or passive) was a
between-group manipulation and type of scenario (shoot, no-shoot) and type of detail (victim, perpetrator, weapon, action, other) were assessed with repeated measures. All effects reported as statistically reliable were significant at or beyond the .05 level.

Participants recalled a larger proportion of details from the no-shoot scenario, $M = .66$, $SD = .123$, than from the shoot scenario, $M = .50$, $SD = .118$, producing a statistically reliable main effect of scenario, $F(1, 38) = 82.15$, $MSE = .0316$, $\eta^2 = .675$. In addition, the main effect of type of detail was statistically reliable, $F(4, 152) = 14.92$, $MSE = .0514$, $\eta^2 = .263$, as was the interaction of scenario and type of detail, $F(4, 152) = 2.95$, $MSE = .0266$, $\eta^2 = .048$. The comparison of recall by active participants and bystander witnesses was not statistically reliable, $F(1, 38) = .40$, $MSE = .116$.

Cell means and standard deviations associated with the interaction of scenario and detail are presented in Table 1.

A Tukey HSD post hoc test of pairwise comparisons of means involved in this interaction indicated that memory of details about the perpetrator in the shoot scenario, $M = .33$, $SD = .094$, was significantly worse than memory of all other types of detail in both scenarios (mean proportion recalled ranged from .51 to .76). Memory of details about the perpetrator in the no-shoot scenario was also inferior to memory of other types of detail but was significantly better than memory of the perpetrator in the shoot scenario. Memory of action did not differ reliably across the two scenarios. In addition, for every type of detail except actions taken by the active officer, police remembered fewer details from the shoot scenario than from the no-shoot scenario.

**DISCUSSION**

The hypothesis that active engagement with an event might influence the accuracy of eyewitness memory was not supported. During actual events, victims and bystanders are unlikely to have the same perspec-

<table>
<thead>
<tr>
<th>Condition</th>
<th>Type of detail recalled</th>
<th>Perpetrator</th>
<th>Victim</th>
<th>Weapon</th>
<th>Action</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td></td>
<td>.33</td>
<td>.53</td>
<td>.54</td>
<td>.54</td>
<td>.58</td>
</tr>
<tr>
<td>$SD$</td>
<td></td>
<td>.094</td>
<td>.103</td>
<td>.204</td>
<td>.286</td>
<td>.326</td>
</tr>
<tr>
<td>No shoot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td></td>
<td>.51</td>
<td>.73</td>
<td>.73</td>
<td>.59</td>
<td>.76</td>
</tr>
<tr>
<td>$SD$</td>
<td></td>
<td>.146</td>
<td>.123</td>
<td>.195</td>
<td>.250</td>
<td>.271</td>
</tr>
</tbody>
</table>
tive, which confounds comparisons of the accuracy of participants and bystanders (Christianson & Hübinette, 1993; Tollerstrup, Turtle, & Yuille, 1994; Yuille et al., 1994). In this experiment, participants were similar both in terms of presumed observation skill (all were police officers) and in viewing perspective. Moreover, although police had greater difficulty recalling details from the shoot scenario than from the no-shoot scenario, this effect did not interact with their participant or bystander status. Thus, it is unlikely that the absence of differences in performance between active participants and bystanders should be attributed to insensitivity of the performance task. This task was sensitive to changes in performance associated with the shoot manipulation.

An additional finding of interest was the observation that officers recalled the fewest details about perpetrators. Memory of the perpetrator also was significantly worse when a shooting occurred. Egith (1993, 1994) noted that weapons attract attention. Loftus, Loftus, and Messo (1987) observed that participants spend more time focusing on a weapon than on other objects in a slide. Therefore, the ability to identify a perpetrator (or other pertinent information) should be impaired when a weapon is present. Usually, weapon focus effects are demonstrated by showing that memory of a perpetrator is worse when a weapon is present than when no weapon is present (Tooley, Brigham, Maass, & Bothwell, 1987).

However, another use of the term weapon focus refers to the tendency to recall weapons better than other types of information (Cutler & Penrod, 1995). For example, Kramer, Buckout, and Eugenio (1990) examined weapon focus effects under conditions in which weapons were visible for either 4 or 12 s. Cutler, Penrod, and Martens (1987) compared the effects of a brandished gun with those of a concealed gun on the accuracy of eyewitness reports. Pickel (1998) used stimulus materials in which a weapon was present in all conditions to evaluate the effects of novelty and threat posed by a weapon on memory accuracy. Thus, weapon focus effects sometimes refer to a pattern of memory for details within an episode in which details about a perpetrator are remembered less accurately than details about a weapon.

To the extent that participants remembered perpetrators less well than weapons in a given scenario, a weapon focus–like effect was obtained in this experiment. Researchers have debated whether weapon focus effects are produced by the narrowing of attention associated with high levels of arousal or by the distracting effects of a novel object (Deffenbacher, 1983, 1994; Egith, 1994; Heuer & Reisberg, 1992; Loftus et al., 1987; Pickel, 1998). In this experiment, although a weapon appeared in both scenarios, the difference between memory of the perpetrator,
$M = 33\%$ correct, and memory of the weapon, $M = 54\%$ correct, was larger in the shoot scenario than in the no-shoot scenario, $M = 51\%$ correct for the perpetrator and $M = 73\%$ correct for the weapon. Thus, although others have demonstrated weapon focus effects under conditions that did not evoke emotional arousal (Heuer & Reisberg, 1992; Kramer et al., 1990; Loftus et al., 1987; Pickel, 1998), this differential recall of weapon and perpetrator detail may be exaggerated when the weapon poses a clear threat. This interpretation must be tempered by the fact that these materials do not include a comparison of memory of the perpetrator when no weapon is present.

Thus, a simulated shooting was found to produce the same harmful effects on memory in law enforcement officers that has been demonstrated in citizen witnesses (Hosch & Bothwell, 1990, Loftus & Burns, 1982; Tollestrup et al., 1994; Yarmey, 1988). This occurred when police were active participants and when they were bystanders. The presumed emotionality and heightened arousal associated with a shooting reduced the ability of law enforcement officers to recall details from the event. However, an integral part of a law enforcement officer’s job is attention to detail and recall of these details during investigation and prosecution of criminal cases. Christianson and his colleagues (Lindholm et al., 1997; Christianson et al., 1998) suggested that police training might contribute to more accurate eyewitness reports from police witnesses than from citizen witnesses. We wanted to determine whether citizens and law enforcement officers differ in their ability to recall details from a traumatic event.

**EXPERIMENT 2**

The major purpose of Experiment 2 was to compare the effects of a simulated shooting on the accuracy of recall by citizen witnesses and law enforcement officers. In addition, the accuracy of memory for different types of information under high and low levels of emotional arousal was examined. Several manipulation checks were added to evaluate whether the shoot versus no-shoot manipulation was successful in creating different levels of stress and physiological arousal in the witnesses. Finally, although we attempted to match the two scenarios used in Experiment 1 in as many ways as possible, the effects of the shoot manipulation might have been confounded with other characteristics of the scenarios. By using a shoot and no-shoot version of each scenario, we hoped to replicate the effects of presumed stress observed in Experiment 1 under methodologically tighter conditions.
METHOD

Participants

The participants were 13 undergraduates enrolled at the University of West Florida (5 men and 8 women) and 16 full-time sworn law enforcement officers with the Pensacola, Florida, Police Department (12 men and 4 women). The mean age of the students was 24 years. The mean age of the police participants was 32 years, with a mean of 8.1 years of experience in law enforcement.2

Materials and apparatus

The scenarios used in Experiment 1 were also used in this experiment. However, the branching capability of the FATS system was used to create shoot and no-shoot versions of each scenario. The shoot and no-shoot versions were identical except for the final 3 s. In the no-shoot version of the domestic disturbance scenario, the driver pulled out a handgun with the barrel pointed downward, immediately put it on the ground, and resumed his stance. The weapon was handled in a nonthreatening manner and officers should not (and did not) shoot during this version. In the shoot version, the perpetrator drew his weapon and immediately aimed and fired at the officer. In the no-shoot version of the attempted abduction scenario, the perpetrator reached toward his pocket but immediately complied when the police participant ordered the perpetrator to raise his hands. (All officers issued appropriate commands at this point in the scenario.) In the shoot version, the perpetrator rapidly pulled a gun from his pocket, aimed, fired once, and either ran off camera or collapsed on the ground (depending on where the shots fired by the participating officer struck). The explicitly threatening use of a gun by the perpetrator made a shoot response appropriate. All police participants fired their training weapons during shoot versions of both scenarios.

Two versions of the questionnaire used in Experiment 1 were created: one for police participants and the other for citizen witnesses. The versions differed only in the perspective assumed for the respondent (e.g., police were asked how often they fired their weapons, citizen witnesses were asked how often the officer fired his or her weapon). In addition, all participants were asked to rate the violence of the scenario and the amount of stress they experienced on 5-point Likert-type scales.

A J & J Standard Volumes system (J & J Enterprises & Garber, 1989) was used to record electrodermal responses (EDRs) as indices of physiological arousal. Disposable finger electrodes were attached to the index and middle fingers of the participant's nondominant hand. Onset of recording was synchronized manually with the onset of the scenario. Mean EDR for successive 2-s bins was recorded for 60 s from the onset of each scenario. Although scenarios ended after either 43 or 52 s, EDR was recorded for a full 60 s to capture any long-latency EDR responses to stress. EDR is a measure of skin conductance in mhos. It is related to galvanic skin response (GSR), which measures skin resistance in ohms. Each measure is the reciprocal of the other. Unlike GSR, EDR values increase with increased levels of arousal (Peek, 1987).
Design

The design was a $2 \times 2 \times 2 \times 5$ mixed factorial design in which participant background (police or citizen) was a between-group comparison, and scenario (domestic disturbance, attempted abduction), shoot condition (shoot, no-shoot), and type of detail recalled (victim, perpetrator, weapon, action, other) were assessed with repeated measures. Two presentation orders were constructed so that both scenarios (domestic disturbance, attempted abduction) were used in shoot and no-shoot conditions across participants. Individual participants viewed one version of each scenario, with groups of participants divided into two presentation orders. Memory data were collected from eight police officers and six students for one presentation sequence (in which the domestic scenario produced the no-shoot condition and the attempted abduction produced the shoot condition) and eight police officers and seven students provided data for a second sequence (in which the attempted abduction produced the no-shoot condition and the domestic scenario produced the shoot condition).

Procedure

Participants were tested in citizen–police pairs (and one in a triad of a police officer and two citizen witnesses). The electrodermal recording procedures were explained and electrodes were applied to each participant (when the triad was run, EDR data were collected from only two participants). The two scenarios were presented with a brief break between scenarios during which EDR recording was completed, EDR data were printed, and equipment was reset for the next scenario. Police participants responded to the action depicted in a scenario as they would during a normal training session. The domestic disturbance and attempted abduction scenarios were presented in a counterbalanced order across participants, but the shoot condition was always presented last. When the second scenario was finished, recording electrodes were removed and participants were given two copies of the questionnaire and asked to complete one copy for each scenario. Participants knew that their memory would be tested at the end of the experiment.

RESULTS

Manipulation checks

EDR data and the two self-report measures (stress experienced and perceived violence) were used to evaluate whether the shoot versus no-shoot manipulation successfully created different levels of stress for participants. Each measure is addressed in turn.

A preliminary analysis indicated that police participants responded with significantly higher EDR values than did citizens, $F(1,42) = 4.32$, $MSE = 1185.442$, $\varepsilon^2 = .072$. Although police participants produced significantly higher EDR values than citizens, participant characteristics did
not interact with any other variables. Mean EDR responses for shoot and no-shoot conditions are presented as functions of time in the scenario for police and citizen witnesses in Figure 1. The same data are presented as functions of time in the scenario for the domestic and abduction scenarios in Figure 2.

These data were rescaled in two ways. Mean EDR was computed for the first 20 s for each participant and used to rescale EDR values. Second, the shooting occurred in the domestic scenario approximately 10 s later than in the abduction scenario. Therefore, the data for the domestic scenario were shifted in time so that EDR responses from the two scenarios were synchronized to the moment when the shooting occurred.

Three important features are depicted in Figures 1 and 2. First, no reliable differences were observed in rescaled baseline EDR between scenarios, shoot conditions, or type of participant. Second, EDR was significantly elevated in both scenarios (and for both police and citizen witnesses) when the action escalated. Finally, EDR underwent an additional and sustained increase when the action in a scenario branched.
to a shooting. In no-shoot scenarios, EDR values gradually declined once the conflict was resolved at the shoot versus no-shoot branch point. These observations were confirmed by an analysis of covariance. Mean EDR for the first 20 s of each scenario was used as a covariate in a $2 \times 2 \times 2 \times 15$ mixed ANCOVA in which participant background (police, citizen) was a between groups manipulation and scenario (domestic disturbance, attempted abduction), shoot condition (shoot, no-shoot), and time in the scenario (in 2-s bins) were repeated measures. There was a statistically reliable effect of time, $F(14, 588) = 15.52$, $MSE = .440$, $\eta^2 = .252$, and a significant interaction of time and condition, $F(14, 588) = 5.02$, $MSE = .440$, $\eta^2 = .085$. This interaction reflects the absence of significant differences in EDR in the baseline data for any of the scenarios but, as expected, significant increases in EDR in both scenarios when the action branched to a shooting. Thus, the analysis of EDR data suggests that both scenarios generated physiological arousal and, more important, that versions of the scenarios that ended in shootings produced significantly more physiological arousal than did versions in which no shooting occurred.

Another way to approach the analysis of the EDR data is to ask how
many participants manifested the predicted pattern of higher arousal when a scenario ended in a shooting. Each participant's EDR function for each scenario was examined to identify participants who responded to the shoot manipulation in the predicted way. Participants were classified as showing the predicted stress response if they produced larger EDR values when the action in a scenario intensified (i.e., over the final 10 s) during the shoot condition than during the no-shoot condition. Of the participants who provided EDR data, 10 police officers (63%) and 7 citizens (58%) responded as predicted. Thus, more police showed the expected elevation in EDR, which is consistent with the self-report data reported next.

Participants reported stress on a 5-point scale in which 5 corresponded to the highest level of stress. Consistent with the EDR data, police reported experiencing significantly more stress, \( M = 3.8 \), than citizen participants, \( M = 2.9 \); \( F(1, 40) = 10.62, MSE = .792, \varepsilon^2 = .19 \). All participants reported that shoot conditions were significantly more stressful, \( M = 3.7 \), than no-shoot conditions, \( M = 3.0 \); \( F(1, 40) = 7.51, MSE = .792, \varepsilon^2 = .137 \). No other main effects or interactions were statistically reliable.

Perceived violence of the scenarios was also rated on a 5-point scale, with 5 representing the highest violence rating. Analysis of the self-report data for perceived violence indicated that participants regarded the shoot conditions as more violent than the no-shoot conditions, \( F(1, 40) = 32.58, MSE = .735, \varepsilon^2 = .435 \). In addition, participants reported that the attempted abduction scenario was significantly more violent than the domestic disturbance scenario, \( F(1, 40) = 5.17, MSE = .735, \varepsilon^2 = .092 \). Police gave both scenarios more extreme violence ratings than did citizen witnesses, \( F(1, 40) = 12.15, MSE = .735, \varepsilon^2 = .214 \). A significant interaction between scenario and condition, \( F(1, 40) = 12.38, MSE = .735, \varepsilon^2 = .217 \), indicated that the shoot versus no-shoot manipulation produced significant differences in perceived violence for the domestic disturbance scenario but not for the attempted abduction scenario. Participants rated the no-shoot version of the domestic disturbance scenario as nonviolent, \( M = 2.1 \), but they gave both the shoot and the no-shoot version of the attempted abduction scenario strong violence ratings, \( M = 4.4 \) and \( M = 3.5 \), that were as extreme as the rating given to the shoot version of the domestic disturbance, \( M = 4.1 \).

**Eyewitness memory performance**

Responses to the memory questionnaire were scored as in Experiment 1 with two exceptions. First, we failed to record verbal commands issued by police participants during data collection, so responses to this question could not be scored and were eliminated from the analysis. Second, the responses to the question about the number of times officers fired
their weapon were scored in two parts. The first part evaluated accuracy about whether the officer fired at all. The second part evaluated accuracy in reporting the number of shots fired. For each cohort of participants, reports of the number of shots fired were compared with the FATS system record of the number of shots fired by the officer in that cohort.

Proportion of correct recall scores were analyzed with a $2 \times 2 \times 2 \times 5$ mixed ANOVA in which participant (police, citizen) was a between-group manipulation, and condition (shoot, no-shoot), scenario (domestic disturbance, attempted abduction), and type of detail (victim, perpetrator, weapon, action, other) were assessed with repeated measures. A significant main effect of condition, $F(1, 48) = 18.75$, $MSE = .036$, $\varepsilon^2 = .266$, indicated that participants recalled a larger proportion of details from scenarios in the no-shoot condition, $M = .75$, $SD = .099$, than in the shoot condition, $M = .65$, $SD = .109$. The significant main effect of scenario, $F(1, 48) = 27.25$, $MSE = .036$, $\varepsilon^2 = .349$, indicated that participants also recalled a larger proportion of details from the domestic disturbance scenario, $M = .76$, $SD = .107$, than from the attempted abduction scenario, $M = .64$, $SD = .087$. A significant main effect for type of detail, $F(4, 192) = 24.10$, $MSE = .025$, $\varepsilon^2 = .320$, combined with significant interactions between details and condition, $F(4, 192) = 10.96$, $MSE = .025$, $\varepsilon^2 = .169$, and details and scenario, $F(4, 192) = 9.58$, $MSE = .025$, $\varepsilon^2 = .149$, indicated that details were not remembered equally well and varied with both shoot condition and scenario. Police and citizen eyewitnesses did not differ reliably in recall accuracy, $F(1, 48) = 1.96$, $MSE = .036$. Note that this comparison also suggests that there were no reliable differences in memory performance by participants and bystanders or by younger and older adults.

The means for proportion correct as a function of type of detail and shoot condition are presented in Figure 3 and proportion correct as a function of type of detail and scenario are presented in Figure 4.

All post hoc analyses reported are Tukey–Kramer HSD tests for unequal sample sizes, computed using the procedures described in Kirk (1995). Analysis of the detail $\times$ condition interaction indicated that in shoot conditions the weapon was remembered better than all other details in that condition. Memory of the perpetrator was reliably worse than recall of actions, but other contrasts were not statistically reliable. In addition, memory of every type of detail in the shoot condition was worse than in the no-shoot condition (although the only statistically reliable contrast was for memory of action). In no-shoot conditions, weapon and action details were remembered better than other details. In addition, memory of the perpetrator was reliably worse than memory of all other types of detail except memory of the victim. All pairwise
comparisons of memory accuracy for action, weapons, and other details were statistically reliable. Among these three types of detail, actions were recalled better than weapon details, which were recalled better than other details.

Post hoc analysis of the detail \times scenario condition indicated that details about the perpetrator, the victim, and other details were more difficult to remember in the abduction scenario than in the domestic scenario. In both scenarios, weapon and action details were remembered better than details about the perpetrator. In the abduction scenario, these details also were remembered better than details about the victim and other details. In the domestic scenario, the only reliable contrasts were between memory of the perpetrator and memory of the weapon, action, and other details. All other contrasts were not statistically reliable.

If increased stress is associated with poorer eyewitness accuracy, it might be possible that a superior ability to recall witnessed information by police was masked by the effects of their heightened stress (as indicated by both self-report and EDR measures). To address this question, analysis of the memory data was repeated with EDR included as a covariate. In this case, mean EDR for the final 10 s of the scenario was used.
as the covariate because this value most accurately represents the amount of stress created by the shoot manipulation. However, this analysis produced results that were identical to those reported earlier, so details of this analysis are not reported here.

**DISCUSSION**

These results provide further evidence of the damaging effect of a shooting on eyewitness memory. Participants correctly recalled more details from no-shoot scenarios than from shoot scenarios when scenario memorability was controlled. Police and citizen witnesses did not differ significantly in the accuracy of their eyewitness reports. Citizens were always bystanders, so these results replicated the findings of Experiment 1, suggesting that bystanders and active participants might not differ in the accuracy of their reports when viewing conditions are controlled. An important difference between the research reported here and previous work comparing the accuracy of victims and bystanders (Christianson & Hübinette, 1993; Hosch & Bothwell, 1990; Tollestrup et al., 1994; Yuille & Cutshall, 1986) is that both bystanders and active participants viewed the scenarios from the same vantage point and had similar ac-
cess to detail information. A different outcome for the comparison of police and citizen witnesses is reported by Lindholm et al. (1997), who also used a video of a simulated crime recorded as it would be seen by a witness. Lindholm et al. found that police recalled more details about the perpetrator than did citizen witnesses, although police and citizens did not differ significantly in their overall level of recall. Their results also differed from those of Experiment 2 in that their police participants reported experiencing less stress than did citizens, whereas police in Experiment 2 experienced more stress than did citizens. However, even when differences in stress were equated across groups by using EDR measures as a covariate, police in Experiment 2 did not differ from citizens in either memory accuracy or in changes in memory accuracy associated with the shoot manipulation.

A second effect that was reliably produced was a weapon focus–like effect. Details about perpetrators were remembered less well than were details about weapons and actions (essentially, memory for an officer’s weapon use). This effect was robust across the two scenarios and the shoot versus no-shoot manipulation. It was obtained even in the no-shoot version of the domestic scenario, in which the weapon was not wielded in a threatening manner. It would be interesting to determine whether the damaging effects of a simulated shooting follow a time course similar to that shown by Loftus and Burns (1982) or Christianson and Nilsson (1984). An analysis of this type would entail precise control over the sequence of events so that the point in time when specific details were encoded could be determined. Because no new details occurred after the shooting in these scenarios and details that appeared before the shooting were available for long durations, such an analysis was not possible for these data.

Finally, recall of information from the abduction scenario was worse, regardless of shoot condition, than recall from the domestic scenario. Both versions of the abduction scenario were rated to be as violent as the shoot version of the domestic scenario. However, the EDR data suggest that the perceived violence of the no-shoot version of the abduction scenario did not generate unusual physiological arousal, so it is unclear whether the difference in recall for the two scenarios should be attributed to unidentified characteristics of the scenarios or to the effects of perceived violence.

**GENERAL DISCUSSION**

Active participants and bystander witnesses did not differ in their ability to recall details about a witnessed event. These results were ob-
EYEWITNESS RECALL

379

tained when bystanders were police officers (Experiment 1) and when they were citizen witnesses (Experiment 2). Researchers who have reported differences in recall by bystanders and participants have expressed concern over the possible confound produced when victims and bystanders differ in their access to information because they have different viewing perspectives (Christianson & Hübinnen, 1993; Hosch & Bothwell, 1990; Tollestrup et al., 1994; Yuille & Cutshall, 1986). In our experiments, participants and bystanders viewed events from the same perspective and recalled details equally well. Thus, vantage point and access to detail information may contribute more to the accuracy of eyewitness memory than the witness’ status as a bystander or participant.

These experiments also showed that overall memory of detail information was lower when a traumatic event (a simulated shooting) occurred, although some details were remembered better than others. More important, we obtained these effects under conditions in which the arousal associated with witnessed events was clearly documented with both a physiological measure (EDR) and self-report measures. In addition, two scenarios that differed in their basic memorability were shown to be rendered less memorable when they ended in a simulated shooting. Experiment 2 demonstrated that the deleterious effect of a weapon on memory of other details did not depend entirely on the perceived threat of the weapon. In fact, the disparity between memory of the weapon and memory of other details was largest in the no-shoot versions of the scenarios. Similarly, Pickel (1998) reported that the unusualness of a weapon is more important than perceived threat in the creation of weapon focus-like effects.

Finally, the comparison of police and citizen eyewitness accuracy continues to produce mixed results. Police and citizens in Experiment 2 did not differ reliably in their ability to recall details from the scenarios. However, Lindholm et al. (1997) found that police remembered a perpetrator better than did citizens. They also found that police identified the weapon in a lineup task more accurately than did citizens. However, police did not differ from citizens in overall memory of events, nor were they more accurate than citizens in a lineup task for the perpetrator. Christianson et al. (1998) presented results that only partially replicate the Lindholm et al. findings. Police obtained the highest scores on all retention measures and differed from one or more comparison groups on memory for a variety of types of information, but their performance did not differ significantly from some of the other groups included for comparison. For example, although police correctly recognized more details about the perpetrator than did teachers, their performance was not significantly different from that of students or police recruits. A similar result was obtained for memory of events.
Moreover, although Lindholm et al. (1997) did not find differences in overall memory, Christianson et al. (1998) did, albeit using different stimulus materials. Christianson et al. obtained significant differences between police and citizen witnesses in recall of information about the victim but did not find reliable differences in recognition for this information. In contrast, police and citizens did not differ in their recall of event details but did differ on the recognition test for these details.

Effect sizes computed for the data from these two studies range from small ($e^2 = .06$ for memory of perpetrators) to moderate ($e^2 = .239$ for memory of peripheral information). Given the difficulty Christianson and his colleagues encountered replicating their results while using a large sample, it is not surprising that we did not find reliable differences between police and citizens based on a smaller sample and possibly more varied stimuli (associated with the manipulation of violence). Comparing the percentage correct scores reported by Christianson and his colleagues for police and student samples with the scores observed for police and citizens in Experiment 2 suggests that a similar (if not statistically reliable) pattern was observed. Citizens correctly recalled less information overall (64%) than did police (76%). Thus, rather than advocate that police are no more accurate than citizen witnesses, we advocate reserving judgment until a sufficiently large body of evidence is available to perform a meta-analysis of this comparison. If the numerical differences between police and citizen accuracy obtained to date eventually emerge as reliable differences, they may prove to be small differences. However, in a criminal investigation or prosecution, correct recollection of a few additional details could sometimes have large consequences for decisions about guilt or innocence.

The observation that police did not differ from citizens in the accuracy of their memory is less surprising than the observation that police experienced more stress than citizen witnesses in Experiment 2 when Christianson and his colleagues report that police experience less stress when viewing violent scenarios. However, police participants in the studies reported here interacted with training materials in the presence of several observers (a citizen witness, the experimenters, and a fellow police sergeant who ran the FATS equipment). Officers who decide to use lethal force know that they will be asked to justify their decision. Moreover, they are sensitive to public concerns about whether the use of lethal force is excessive. Police participants might have experienced additional stress associated with the perception that their actions and judgment might be evaluated.

Finally, considerable concern has been raised about applying eyewitness research to the evaluation of specific eyewitness testimony (Beke- rian, 1993; Christianson, 1992; Cutler & Penrod, 1995; Sporer et al.,
Although some researchers have focused their concern on individual differences among witnesses, equal concern should be directed toward the issue of generality across experimental materials. Researchers have relied on a small set of stimulus materials to evaluate eyewitness memory, often replicating effects across experiments using known materials. There is merit in replicating effects while controlling for stimulus materials, but it is also important to know that these effects will generalize beyond these materials. We took considerable pains to select two scenarios that were comparable in duration, complexity, and other factors that we thought might influence memorability. However, the variability in our outcomes clearly suggests that seemingly minor differences in stimulus materials could have important consequences for the accuracy of eyewitness memory. Thus, research using a greater variety of stimulus materials is needed.

**Appendix. Questionnaire and instructions to participants**

You have just had the opportunity to interact with a scenario which is similar to a situation you may actually encounter in your career. If this had been a real incident, you would be expected to complete a report containing the details of this incident as part of your duties. Your report would contain sufficient details to document what happened, justify your actions, help identify a suspect, or prepare you and others for court.

The following is a list of questions designed to test your recall for details of this scenario. You should provide the same information you would provide in an actual report. You are asked to respond to each question. Each question will require that you describe details. Please provide as much specific information as you can recall. If the you believe the question does not apply to this scenario, please indicate "N/A" in the answer space. Since the answers to these questions are dependent upon your recall, please do not discuss the questions with other participants.

1. Describe the victim's race.
2. Describe the victim's sex.
3. Describe the length of the victim's hair.
4. Describe the color of the victim's hair.
5. Describe the victim's facial hair.
6. Describe the victim's hat.
7. Describe the victim's jacket.
8. Describe the victim's shirt/blouse.
9. Describe the victim's pants/ skirt/dress.
10. Describe the suspect's race.
11. Describe the suspect's sex.
12. Describe the length of the suspect's hair.
13. Describe the color of the suspect's hair.
14. Describe the suspect's facial hair.
15. Describe the suspect's hat.
16. Describe the suspect's jacket.
17. Describe the suspect's shirt/blouse.
18. Describe the suspect's pants/skirt/dress.
19. Describe the suspect's first weapon.
20. Describe the suspect's second weapon.
21. In which hand was the suspect holding the first weapon?
22. In which hand was the suspect holding the second weapon?
23. What was the suspect doing with the first weapon when you initially observed it?
24. What was the suspect doing with the second weapon when you initially observed it?
25. What was the make of the vehicle at which this incident occurred?
26. What was the body style of the vehicle at which this incident occurred?
27. What was the color of the vehicle at which this incident occurred?
28. At which side of the vehicle did the incident occur?
29. What verbal commands did you issue?
30. How many times did you fire your weapon?

Notes

The authors acknowledge the assistance of Chief Louis Goss, Karen Castille, Captain George Underwood, Lieutenant John Mathis, Sergeant Greg Moody, Officer Jim Enterkin, Officer Waylon Mann, Officer Chris Wilkinson, and the men and women of the Pensacola Police Department for use of the FATS system and access to volunteers. Experiment 1 was a master's thesis submitted by Thomas Johnson to the University of West Florida. These data were presented at the 1994 meeting of the Southeastern Psychological Association. Ginger Baker and Gabrielle Whipple assisted with data collection for Experiment 2. Frank Andrasik provided advice and equipment for recording EDRs. These data were presented at the meetings of the 1997 Psychonomic Society and the 1998 Florida Conference on Sensation, Perception, Cognition, and Action. Finally, the authors want to thank Frank Andrasik, Kenneth Deffenbacher, R. Edward Geiselman, and anonymous reviewers for comments and suggestions on two earlier drafts of this manuscript.

Correspondence about this article should be sent to Claudia J. Stanny, Psychology Department, University of West Florida, 11000 University Parkway, Pensacola, FL 32514 (e-mail: cstanny@uwf.edu). Thomas C. Johnson is now at Mississippi State University. Received for publication June 5, 1998; revision received January 31, 1999.

1. Effect sizes were computed using epsilon squared, as discussed by Keppel (1982).
2. Questions about the possible effects of differences between these groups based on age or gender of participants are not readily answered in the litera-
tured. Although older adults (age 60 or older) sometimes are reported to recall witnessed events less accurately than young adults, these differences are not obtained consistently (Bornstein, 1995; Deffenbacher, 1991). Moreover, with few exceptions (e.g., Yarmey & Yarmey, 1997), young adults are defined in such a way that most participants in both groups in the current experiment would fall in the young adult range. Yarmey and Yarmey (1997) reported that men and women were similar in their ability to recall characteristics of a person briefly encountered. They also reported one of the few findings of a negative relationship between age and recall over a broad range of ages. However, it is unclear from their data how large a difference in memory performance we can expect based on an average age difference of only 8 years.

3. The decision to rescale the EDR data was based on the assumption that changes from baseline are more important as a measure of stress than are raw values for EDR. Individual differences in baseline EDR can be large. Initial analysis of raw EDR values produced an interaction of participant type, scenario, and condition that was driven by individual differences in baseline EDR. Mean EDR for the first 20 s of the recording period was chosen for rescaling because visual inspection of the unscaled group data showed that there were no changes in baseline EDR until well after 20 s. Thus, the first 20 s of EDR data were used for rescaling and the remaining data were used to determine that EDR did not differ between scenarios or conditions until the shooting occurred. The only effect that rescaling had on the analysis of EDR data was to eliminate effects created by individual differences in baseline.

4. A smaller number of levels of time in the scenario appears in this analysis because EDR data from some time bins in one scenario no longer had corresponding data in the second scenario when the scenarios were synchronized to time of the shooting. These unmatched EDR measures were discarded. In addition, not all participants provided usable EDR data. Two police officers and one citizen participant did not contribute EDR data. Data from one officer and the citizen participant were lost because of a recording error. The second officer was not wired for EDR to allow data collection from two citizen participants.

References


