Law and Human Behavior
2016, Vol. 40, No. 1, 23–35

0147-7307/16/$12.00 http://dx.doi.org/10.1037/lhb0000159

Memory and the Operational Witness: Police Officer Recall of Firearms Encounters as a Function of Active Response Role

Lorraine Hope
University of Portsmouth

David Blocksidge
Metropolitan Police, London, United Kingdom

Fiona Gabbert
Goldsmiths, University of London

James D. Sauer
University of Tasmania

William Lewinski
Force Science, Mankato, Minnesota

Arta Mirashi
University of Portsmouth

Emel Atuk
University of Portsmouth

Investigations after critical events often depend on accurate and detailed recall accounts from operational witnesses (e.g., law enforcement officers, military personnel, and emergency responders). However, the challenging, and often stressful, nature of such events, together with the cognitive demands imposed on operational witnesses as a function of their active role, may impair subsequent recall. We compared the recall performance of operational *active* witnesses with that of nonoperational *observer* witnesses for a challenging simulated scenario involving an armed perpetrator. Seventy-six police officers participated in pairs. In each pair, 1 officer (active witness) was armed and instructed to respond to the scenario as they would in an operational setting, while the other (observer witness) was instructed to simply observe the scenario. All officers then completed free reports and responded to closed questions. Active witnesses showed a pattern of heart rate activity consistent with an increased stress response during the event, and subsequently reported significantly fewer correct details about the critical phase of the scenario. The level of stress experienced during the scenario mediated the effect of officer role on memory performance. Across the sample, almost one-fifth of officers reported that the perpetrator had pointed a weapon at them although the weapon had remained in the waistband of the perpetrator's trousers throughout the critical phase of the encounter. These findings highlight the need for investigator awareness of both the impact of operational involvement and stress-related effects on memory for ostensibly salient details, and reflect the importance of careful and ethical information elicitation techniques.

Keywords: eyewitness memory, stress, arousal, law enforcement, interviewing

Law enforcement officers, military personnel, and others in civil or emergency response occupations are frequently involved in dynamic, challenging incidents. Depending on their particular operational mandate, these "operational witnesses" may need to act to preserve life, protect citizens, neutralize threats, initiate recovery, or engage in some combination of related activities to resolve an incident. Accurate and detailed accounts of the incident, and information about the operational witness's own activities and that of colleagues, may be important for subsequent investigations and the eventual delivery of justice (Alpert, 2009). Nowhere is this more critical than in the case of shootings by armed police officers who are authorized by the State to discharge a weapon in the course of their duty to protect and avert imminent threats to life (ACPO [Association of Chief Police Officers], 2011; Armed Policing Authorised Professional Practice, 2013). The current research examined the effects of active involvement on eyewitness recall memory. We compared the recall performance of operationally active witnesses responding to a (simulated) threatening incident with that of nonoperational witnesses, or bystanders, to the same event; hereafter referred to as active and observer witnesses, respectively.

Recent high profile cases, such as the shooting of Mark Duggan in the United Kingdom, serve to highlight tensions in the investigation of armed incidents. Duggan was shot by an armed officer in London. A Public Inquest, conducted between September 2013

This article was published Online First October 5, 2015.

Lorraine Hope, Department of Psychology, University of Portsmouth; David Blocksidge, Metropolitan Police, London, United Kingdom; Fiona Gabbert, Department of Psychology, Goldsmiths, University of London; James D. Sauer, School of Psychology, University of Tasmania; William Lewinski, Force Science, Mankato, Minnesota; Arta Mirashi and Emel Atuk, Department of Psychology, University of Portsmouth.

Correspondence concerning this article should be addressed to Lorraine Hope, Department of Psychology, University of Portsmouth, Portsmouth PO1 2DY, United Kingdom. E-mail: lorraine.hope@port.ac.uk

and January 2014 concluded that the shooting had constituted a lawful killing (see http://dugganinquest.independent.gov.uk/latest-news.htm). However, in the course of both the investigation and inquest, questions were raised with respect to officers' accounts of the incident, and perceived inconsistencies in the statement provided by the officer who discharged his weapon were widely reported in the media. Perhaps unsurprisingly, skepticism over the accounts provided by officers in the aftermath of fatal shootings is well documented both in the U.K. and internationally, and generally focuses on the potential for police collusion or corruption (Braidwood, 2010; Heaton-Armstrong & Wolchover, 1993, 2009; see also Hope, Gabbert, & Fraser, 2013). Although the deliberate fabrication of evidence lies beyond the scope of the current article, there is a less controversial explanation for at least some of the apparent inconsistencies and inaccuracies in the "honestly held" accounts provided by operational responders in the aftermath of stressful incidents. The common lay belief that memory operates like a video-recorder—often subscribed to by investigators, legal experts, and potential jury members—is woefully inaccurate (Simons & Chabris, 2011; see also Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006). Memory is a dynamic and reconstructive process, susceptible to error and distortion (Schacter, 1999). It is well documented, in over 40 years of research, that eyewitness memory is fallible—even under optimal encoding conditions (Conway, 2012).

The recall of operationally active witnesses may suffer additional performance decrements because of their role in challenging response contexts; for instance, as a result of increased cognitive demands associated with response generation or decision-making. Understanding how memory performs under such conditions is important for a number of reasons. First, there is a dearth of research examining memory performance under realistic operational conditions (simulated or otherwise). Second, it is important that procedures and policies relating to the treatment of operational witnesses by independent investigators or other agencies are (a) well-informed with respect to the malleability of memory and (b) appreciate the necessity for careful and ethical memory elicitation practices. Third, it is critical that practitioners in the legal system, whether lawyers, prosecutors, or judges, are cognizant of the potential for naturally occurring memory errors, gaps, and inconsistencies among all witnesses but particularly operational witnesses, such as armed officers, who are often expected to provide detailed and accurate accounts of challenging incidents they were immersed in.

In light of these concerns, we examined the memory performance of witnesses with operational duties (i.e., police officers actively responding to a simulated incident) with that of witnesses who had not been operationally deployed (i.e., observers). We also explored the effect of schema-driven expectation in operational contexts on memory error.

Stress, Arousal, and Memory Performance

Armed officers regularly find themselves in unpredictable, dangerous environments and may experience varying degrees of emotional arousal and/or stress response (Meyerhoff et al., 2004). Research on the effect of arousal on cognitive processes in applied training settings reveals the cognitive and memory difficulties experienced in high-stress environments. For example, Morgan and colleagues (2004) tested the memories of soldier participants who had been exposed to high levels of interrogation stress, including physical confrontation, in an intensive survival school training exercise. Memory performance, in terms of the recognition of a target individual who had physically confronted and threatened them for over 30 min, was impaired after high stress (vs. low stress) interrogations. More recently in a study involving 861 soldiers in a similar survival training simulation, Morgan, Southwick, Steffian, Hazlett, and Loftus (2013) observed that memories for stressful events, like memories for more mundane events, are susceptible to misleading postevent information. Challenging field environments are also associated with significant impairment in selective and sustained attention (Leach & Ansell, 2008) and reduced working memory capacity (Leach & Griffith, 2008). Focusing on the performance of police officers in simulated operational settings, Hope et al. (2012) examined the effects of physiological stress, as a function of exertion, on recall and recognition and found that police officers who had been exerted reported significantly fewer correct details about an encounter, and were significantly less likely than nonexerted officers to identify an encountered target individual.

The precise mechanisms underpinning memory impairment as a result of stress and arousal are difficult to directly delineate in the applied context. A large body of literature confirms that emotionally arousing events are remembered better than neutral events (e.g., Payne et al., 2006) with neurobiological research, in particular, suggesting that stress hormones can enhance memory consolidation (McGaugh, 2013; Roozendaal, 2000). Researchers have also speculated that attentional narrowing under arousal underpins this recall advantage such that memory for central or important stimuli is enhanced (on the grounds that such items "capture attention") while memory for peripheral items is impaired (Safer, Christianson, Autry, & Österlund, 1998; see also Easterbrook, 1959). However, there are a number of problems applying this rather simplistic account to the complex interaction between stress and arousal on memory in applied contexts. First, the detrimental effects of high levels of stress experienced in naturalistic settings have been well-documented (e.g., Morgan et al., 2004, 2013). In fact, in their meta-analysis, Deffenbacher et al. (2004) identify what they describe as a "catastrophic" decline in memory performance at higher stress levels. As such, the effect of arousal on memory performance reflects an inverted U-shaped curve with memory for events best when stress levels are moderate (Morley & Farr, 2012). Thus, while arousal may activate the amygdala (Adolphs, Tranel, & Buchanan, 2005; Phelps, 2006), higher levels of stress work to disrupt hippocampus function, impairing memory for sensory detail and visuospatial working memory (Shackman et al., 2006; for an extended version of this argument see Davis & Loftus, 2009). Furthermore, pharmacological research observes that stress hormones in the form of glococoricoids and catecholamines (adrenaline/nor-adrenaline), naturally released during stress (De Kloet et al., 1998), have variable effects on memory, depending on a number of modulatory factors (Lupien & Lepage, 2001; Wolf, 2003). In particular, the release of cortisol (or its administration in placebo-controlled pharmacological studies) is associated with impaired memory retrieval (deq Uervain et al., 2000; Kuhlmann, Piel, & Wolf, 2005; Wolf et al., 2001). In a sample of Special Forces candidates evaluated in the course of an intense naturalistic stressor, high levels of cortisol secretion were associated with impaired cognitive performance relative to a control group (Taverniers, Taylor, & Smeets, 2013).

A second problem with "attentional narrowing" accounts, as noted by Davis and Loftus (2009), is that emotional arousal may not narrow attention in all cases, and certainly not to predictable stimuli. In fact, stress impairs executive function (Schoofs, Wolf, & Smeets, 2009), including the ability to control attention or where it is directed (Banks, Tartar, & Welhaf, 2014). Laboratory work demonstrating supposed attentional narrowing typically uses relatively uncomplicated or unambiguous stimuli (e.g., the deliberate presentation of a single weapon in much of the so-called weapon focus literature) in a third-hand presentation format (e.g., video or slides). In real life situations, there are likely to be many conflicting draws on attention as a function of (a) a more complicated interactive scene and (b) the need to respond to an incident, which may further deplete cognitive resources. Taken together, these factors make it difficult to predict what will be remembered from stressful naturalistic events. Thus, the first aim of the current study was to examine the effects of arousal, experienced in a challenging yet controlled event, on officer recall as a function of response role.

Attentional Load, Training, and Expectations

Both operationally active and lay (or civilian) witnesses' memories for an incident may be impaired as a consequence of high stress levels during encoding. However, active witnesses have the additional task of deriving an appropriate response strategy (in light of various and potentially transient contextual factors, and taking into account their own safety and that of others in the vicinity), planning the effective execution of that strategy and then taking action accordingly (see Eyre & Alison, 2007). Obviously, an important contextual factor in minimizing the cognitive "drain" of such activities should be the effective training of operationally active witnesses such that response options are fluently available (e.g., recognition-primed decision-making, Klein, 1998; see also Clark, 2008; Healy, Kole, & Bourne, 2014). However, by its nature, training might set expectations about the likely outcomes of particular scenarios. According to Neisser's (1976) notion of the "perceptual cycle," individuals perceive, interpret, and revise their understanding of information using both top-down and bottom-up processes. Certain stimulus properties elicit an attentional response and, inevitably, expectations derived from existing schemas guide further interpretation and recollection (e.g., Tuckey & Brewer, 2003; see also Most, Scholl, Clifford, & Simons, 2005). For example, when specific schemas are activated, visual attention is likely to be directed toward schema-relevant items (Eberhardt, Goff, Purdie, & Davies, 2004). Evaluating an emerging situation and responding appropriately are critical police activities, particularly for high stakes incidents. However, these activities will likely take place under conditions of time pressure, depleted cognitive resources and additional cognitive load—conditions under which such evaluations, and subsequent recollection, may be particularly vulnerable to schema-driven or expectation-based errors (e.g., Kleider, Pezdek, Goldinger, & Kirk, 2008; Sherman & Bessenoff, 1999; Tuckey & Brewer, 2003; see Betts & Hinsz, 2013). If, under such circumstances, the available schema posits that Action X is usually followed by Action Y, then "interpretations or classifications made on the basis of emotion- or expectation-weakened identification criteria, unchecked by disconfirming evidence, enter long-term memory uncorrected and become the basis of distorted witness reports" (Davis & Loftus, 2009, p. 182). Thus, the second aim of the current study was to examine whether active witnesses, like other professionals in high reliability roles (e.g., Plant & Stanton, 2012), may be vulnerable to expectation-based errors that impact the accuracy of their subsequent accounts.

Current Study

Our research was motivated by two main research questions. First, we examined whether there were differences in the accounts provided by operational (active) and nonoperational (observer) witnesses where both had been exposed to the same incident. To achieve this we put pairs of police officers into an immersive and stressful simulated scenario where one officer was instructed to respond as they would usually in the course of their duty (i.e., an active operational witness) while the other was instructed to simply observe the scenario. To our knowledge, this is the first study to adopt a methodology of this kind to examine the effects of active role involvement in an eyewitness context. We predicted that officers allocated to the active response role would show an increased physiological response in the scenario, reflecting by proxy, increased stress as a function of their active response role. On the grounds of this increased arousal, and consistent with the theoretical accounts outlined above, we predicted that the quality and quantity of free recall reported by the active response officers would be impaired relative to the observers.

In addition to the free-recall task, both witnesses were asked a series of detailed closed questions. These questions were included to contribute data and inform current practice and policy in the investigation of shooting incidents. In a number of recent cases, after providing written statements, operational witnesses have been asked to respond to long lists of additional, detailed closed questions (see Dodd & Travis, 2014). Research on investigative interviewing has long documented the problems associated with closed question interviewing approaches (cf. open questions), not least the dangers of leading, closed questions (e.g., Sharman & Powell, 2012). Current "gold standard" interview techniques, such as the Cognitive Interview (Fisher & Geiselman, 1992) and the National Institute for Child Health and Development Protocol (NICHD; Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007) endorse the use of open-ended questions with compliant witnesses (for overview, see Vrij, Hope, & Fisher, 2014). In the current study, participants were exposed to an extended set of closed questions modeled on the question style often adopted by investigators (see Method). In addition to documenting accuracy rates for such questions, we were particularly interested in the recall performance of active witnesses (relative to their observing cowitnesses) on questions pertaining to use and location of the target's weapon as such questions are, unsurprisingly, a central focus of investigations following police shootings. Specifically, we predicted that additional demands on the resources of operational witnesses in conjunction with higher arousal and stress levels, both likely to occur when police officers were required to discharge (or consider discharging) their own weapon, would impair recall for such information.

Our second independent research question concerned the possible effect of expectations as they relate to memory performance under challenging operational conditions. Deliberately, the scenario was designed to trigger schema-driven expectations regarding the likely action of the perpetrator. In the final sequence of the scenario, the perpetrator who was at this point known to be armed, turned quickly to face the officers, throwing out his hands in front of him. However, the weapon (a gun) remained in the waistband of his trousers. In light of the memory deficits reported in previous research (e.g., Morgan et al., 2013) and well-documented effects of schema-reliance (e.g., Kruglanski & Freund, 1983; Tuckey & Brewer, 2003), we predicted that memory reports provided by officers in the Active (cf. Observer) condition may be particularly vulnerable to the expectation-based error that the perpetrator would point the weapon at them. For the current study, it should be noted that our research questions were entirely focused on postevent recollection by officers and not behavioral outcomes, such as shoot/no-shoot decisions that are explored elsewhere (e.g., Akinola & Mendes, 2012; Nieuwenhuys, Savelsbergh, & Oudejans, 2012).

Method

Participants

Eighty-seven serving Canadian law enforcement officers affiliated to a metropolitan force were recruited. Because of technical difficulties, incomplete data or single participant sessions (due to unavoidable no-shows), the final sample comprised 76 participants. Participants (64 males) were aged 22 to 59 years of age ($M = 37$ years, $SD = 7.99$). Most of the sample was at Constable rank (94%) with the remainder at Detective (3%) and Sergeant (3%) rank. Recruited officers represented a range of experience with an average length of service of 147.75 months of service ($M = 12.31$ years, $SD = 89.47$ months).

The purpose or nature of the study was not revealed in advance—participants were led to believe that they were taking part in research related to officer response. Officer participation in the research was voluntary and took place during work hours at prearranged times with the full agreement of shift supervisors. Although the research was organized in collaboration with the training division of the force, test sessions did not constitute formal training events. Officers were not paid for their participation and received no work-related rewards for taking part. In addition to adhering to standard ethical principles and considerations, detailed Informed Consent procedures assured officers that their individual professional performance was not being assessed and reiterated confidentiality procedures.

Design

Officers were randomly assigned to either the "Active Officer" or "Observer Officer" condition during the encoding phase. Conditions were paired such that each Active Officer viewed the scenario with an Observer Officer. All participants completed the same test materials. The experimental data were collected over a 5-day period. In a typical test day, eight pairs took part in live scenarios.

Materials

Briefing video. Before deployment into the main scenario, all officers viewed a short "briefing" video that depicted the initial hostage-taking incident. It showed students taking part in a classroom-based seminar with their professor and was filmed on a cell phone (as if from the perspective of one of the students). The perpetrator (an apparently disgruntled student) entered the classroom and engaged the professor in a discussion about poor grades. The perpetrator became increasingly agitated, drew a knife from his pocket, and took the professor and a student hostage. Toward the end of the video, the other students are seen rushing from the room, initiating calls to the police. The film, recorded using an iPhone, lasted 2 min 10 sec and was of high quality with clear audio.

Experimental scenario. The scenario was developed in the course of extensive discussion with firearms instructors and police trainers. Three key objectives guided the scenario development: the need for (a) a relatively complex scenario which the participants could be questioned about in detail, (b) a scenario with challenging elements that would produce a natural physiological stress response, and (c) a realistic scenario officers might encounter in the course of their duty and be reasonably expected to respond to.

All officers encountered the same "augmented reality" scenario, lasting 4 min, which incorporated prerecorded and live elements. The prerecorded elements of the incident were presented as "live" CCTV footage and were integrated with fully scripted, tightly controlled live elements, reenacted for each pair of participants, using three actors (one male "perpetrator," two male "hostages"). At the outset, the perpetrator was shown via a CCTV-feed, threatening the two hostages in a hallway (outside the classroom where they had originally been located in the briefing video). One of the hostages was then released and could be seen walking down a hallway, appearing first in the "CCTV footage" and then in reality through a window in the classroom through to the same hallway. This visual device was incorporated to fully establish the link between the apparent CCTV footage and elements of the scenario enacted in real-time activities. This method appeared successful as, during debriefing, a number of officers indicated surprise that any elements had been prerecorded and all officers reported the scenario as a single integrated event. Shifting from prerecorded footage to live interaction, the perpetrator then entered the classroom (where the officer participants were located) using a hostage as a shield and holding a knife to the hostage's neck. After issuing various demands, he set the hostage free and threw the knife to the ground before retreating to the hallway and closing the door to the classroom. He could then be seen, on the CCTV, tucking the gun into the waistband of his jeans. In the final live interaction, the perpetrator reentered the classroom in an agitated manner. The gun remained in the waistband of his trousers throughout.

Recall tasks. The recall tasks comprised two different response formats—Free Recall and Closed Questions. In the Free Recall task, which was presented first, there were two sections. The first section requested details of the briefing information encountered at the outset (Briefing Phase). The second section requested details of the main scenario involving the live CCTV footage and perpetrator (Response Phase). Instructions at the start of both sections asked participants to report as much information as they could remember about the event, emphasizing that their account should be "as complete and accurate as possible." Participants were also instructed against guessing. In the Closed Questions task, participants answered 94 questions adapted, in terms of style and content, from the type of questions posed by external investigators in such circumstances (as mentioned by Dodd & Travis, 2014). Ninety-one of these questions sought factual and verifiable information about the incident, three questions asked for a more subjective assessment of the perpetrator (e.g., "What was the demeanor of the perpetrator?" and "Describe his facial expression"). A subset of these closed questions (14 questions; *target questions*) were identified by legal and police training advisors as important questions from an investigative perspective with respect to an officer's recall of the critical response phase. These question subsets were categorized as "Perpetrator Action" (e.g., *Did he turn to the left or to the right? What position did he move his arms to?*), "Officer Response" (e.g., *What action(s) did the other officer(s) in the room take at this point? How many shots were fired and by whom?*), and "Perpetrator Weapon" (e.g., *Did the culprit discharge a weapon during the incident? If yes, how many shots did he fire? Where was the gun at the end of the scenario?*).

Procedure

Participants were recruited in pairs and were randomly allocated to the role of Active Officer or Observer Officer by virtue of their choice of seat in the waiting area. Both officers were fitted with Polar Heart Rate monitoring belts, equipped with safety glasses, and given general instructions about their role. Active Officers were instructed to respond to the scenario event as they would normally in the course of their duty. Observer Officers were instructed to take no active response role and to simply observe what happened during the scenario. A verbal briefing by a "senior officer" informed them that there was an ongoing hostage situation involving a male perpetrator armed with a knife in a remote corridor area. They were told that negotiations had been underway but that the perpetrator had stopped communicating in the past 30 min. The Active Officer was provided with a training handgun loaded with five blank rounds (i.e., their weapon was available for discharge) and informed s/he was part of an initial response team with the objective of moving forward into the classroom adjacent to the corridor containing the hostages to tactically assess the situation and intervene, or advise other teams available to intervene as necessary.

Before entering the critical response phase, both participants, seated side by side, viewed the briefing video on a laptop screen. They were told that this was "cell phone footage that a witness captured as the situation developed this morning." This concluded the briefing phase. At the outset of the critical response phase, both officers were then taken into the classroom where they could view the live CCTV footage of the ongoing incident. Both were equipped with headsets to ensure they could hear the footage soundtrack and also to mitigate any effects of any verbalization by their coparticipant (some participants initiated verbal comments with respect to the evolving scenario; we noted no apparent effect of this verbalization on reporting and observed that as the scenario escalated, verbal commentary typically ceased). A confederate officer initiated the footage when they entered the room and ensured both participants stood in preallocated side-by-side containment positions (behind a desk) that shared an equally clear view of both the CCTV footage (within 2 feet) and live action space (within 15 feet). A senior police instructor monitored each trial from a health and safety perspective and ended each trial with a whistle blast. Each trial was captured on two digital cameras.

There was a delay of 45–50 min between the end of the response phase and the recall tasks. The purpose of this delay was twofold. First, to allow attenuation of any immediate stress response and second, to mimic a minimum delay before an initial interview with an investigator. During this time, officers were seated at separate desks in silence. No discussion with their partner (or anyone else) was permitted and this instruction was closely monitored by the researcher. After the delay, participants completed the recall tasks. When they had completed the free recall component, they were given the closed questions and worked through them in sequential order. Drawing on an approach devised by Scoboria and colleagues (e.g., Scoboria & Fisico, 2013) to better understand the meaning of "Do Not Know" responses in interview contexts, participants were then asked to clarify any "Do Not Know"/"Unknown" responses by selecting one of four options for each response of that type: (a) I did not report an answer because the information was not present in the event (*no one could answer this question*); (b) I did not report an answer because I could not recall the specific information from the event (*someone else might be able to answer this question, but I cannot*); (c) I did not report an answer because I truly do not remember (*I do not know whether it is possible to answer this question or not from the information in the event*); or (d) I did not report an answer because I was not confident enough that it was correct (*I could provide an answer if pushed, but it might be wrong*).

No time limits were imposed and participants took, on average, 2 hr to complete the recall tasks and final classification task. Officers were then fully debriefed and thanked for their contribution to the research.

Coding

Both free recall sections (Briefing Phase and Response Phase) were coded for quantity and accuracy. Using a coding scheme adapted from Hope, Gabbert, and Fraser (2013), each unit of information reported was categorized as either correct or incorrect. A second coder, blind to experimental condition, coded a random sample (15%) of the free recall reports. Intercoder reliability for briefing phase recall was $\kappa = .72$, $p < .001$, 95% confidence interval (CI) [0.47, 0.98] and for response phase recall was $\kappa = .81$, $p < .001$, 95% CI [0.58, 1.00], suggesting substantial to high levels of agreement (Landis & Koch, 1977). Accuracy rate was calculated by dividing the total correct items by total responses (correct and incorrect).

For the 94 closed questions, model correct answers were agreed by the research team in a detailed review of the scenario videos. No divergence occurred during any trial that would have resulted in a different possible answer for any question. The closed question responses were coded as either correct or incorrect. "I do not know" (and variations thereof e.g., "unknown," "not sure") responses were also recorded. Responses to questions concerning the demeanor of the perpetrator elicited subjective responses ("*he looked angry*"), and consistent with coding conventions in previous research, were not included in the analysis of recall data.

Results

Individual officers took part in pairs (in this case, distinguishable pairs; Kenny, Kashy, & Cook, 2006) such that all participants were exposed to the same highly controlled scripted and videoed stimuli. Given this dyadic design structure, it was necessary to assess the degree of nonindependence before proceeding with the main analyses. Following Kenny et al.'s (2006) approach, we conducted a preliminary analysis to establish the extent to which responses within the pairs were correlated. First, the dataset was restructured to apply SPSS (Chicago, IL) syntax, developed by Alferes and Kenny (2009, see Supplementary Materials), to compute the Pearson product–moment correlation coefficient between dyad members and perform a t test of the null hypothesis that the population correlation is 0. For the free recall data for the 38 intact pairs in the dataset, $r(38) = .12$ [−.21, .42], $t(36) = .74$, $p = .47$. The same analysis was conducted for the cued recall data. The correlation and associated t test were also nonsignificant; $r(38) = .21$ [−.12, .50], $t(36) = 1.28$, $p = .21$. Given independence between dyad member scores for both recall tasks, each individual was used as the unit of analysis in subsequent analyses; see Kenny et al. (2006) for informative discussion around measuring (non)independence in dyadic data.

Physiological Response

Usable heart rate recordings were available for 61 participants in the sample (33 participants in the Active Officer condition and 28 participants in the Observer Officer condition). The mean resting heart rate (HR) recorded over a 10 min period 1 hr after the critical response phase was 68.01 beats per minute (bpm; $SD = 9.17$) and was roughly equivalent between groups (Active Officers $M = 68.81$, $SD = 8.68$; Observer Officers $M = 67.13$, $SD = 9.80$). HR measurements recorded during the response phase showed a range of physiological arousal response with a range of 76–164 bpm (see Table 1). The average maximum HR (mHR) recorded during the response phase were significantly higher for Active Officers than Observer Officers, $t(59) = 2.89$, $p = .005$, $d = 0.75$, 95% CI [0.22, 1.28]. HR variability (HRV; Thayer, Ahs, Fredrikson, Sollers, & Wager, 2012), another measure of workload under stress measured over a 1 min period during the response phase, also differed between groups, $t(59) = -2.30$, $p = .025$, $d = 0.59$, 95% CI [0.06, 1.11]. Lower HRV is associated with increased stress (see meta-analysis by Thayer et al., 2012).

Recall Performance

Free recall. For the free recall data, the dependent variables of interest were quality (as reflected in the accuracy of accounts) and quantity (as reflected in the amount of information provided). As quantity is most usefully examined in terms of the amount of correct and incorrect items, it should be noted that the quality and quantity measures are not independent.

Recall of briefing phase. There was no significant difference between conditions for the amount of correct, $t(74) = 0.06$, $p = .96$, $d = 0.01$, 95% CI [−0.44, 0.47] or incorrect, $t(74) = 0.15$, $p = .88$, $d = 0.03$, 95% CI [−0.42, 0.49] information reported about the events viewed in the briefing video. Similarly, there was no difference between conditions in the overall accuracy rate for this information, $t(74) = -0.35$, $p = .73$, 95% CI [−0.04, 0.02], $d = 0.13$, 95% CI [−0.58, 0.32]; see Table 2.

Recall of response phase. There was a difference between conditions for the number of correct details reported about the critical response phase, such that Active Officers reported significantly fewer correct details than Observer Officers (see Table 2), $t(74) = -2.74$, $p = .008$, $d = 0.63$, 95% CI [0.15, 1.09]. There was no difference in the amount of incorrect information reported between the experimental groups, $t(74) = 0.87$, $p = .39$, $d = 0.20$, 95% CI [−0.25, 0.66]. The overall accuracy rate did not differ between the groups either, $t(61) = -1.46$, $p = .15$, $d = 0.32$, 95% CI [−0.77, 0.14].

Integrating Physiological Response and Free Recall Performance

The analyses suggest a link (a) between the role the officer was assigned to (active vs. observer) and the degree of arousal experienced (using maximum HR as a proxy measure for the peak of that arousal) and (b) between role and free recall performance, specifically the amount of correct information reported about the response phase. These associations might be formulated as $X \rightarrow M$ and $X \rightarrow Y$, respectively, in terms of the meditational model $X \rightarrow M \rightarrow Y$ (where X is the independent variable, Y is the dependent variable and M is the mediating variable; Baron & Kenny, 1986). In line with Baron and Kenny's (1986) recommendations for establishing mediation, we constructed three regression equations. First, regressing mHR on experimental condition ($X \rightarrow M$) was statistically significant, $\beta = -15.72$, 95% CI [−26.77, −4.62], $t(59) = -2.89$, $p = .005$. Second, regressing correct response phase free recall performance on experimental condition ($X \rightarrow Y$) was significant, $\beta = 7.17$, 95% CI [2.17, 12.34], $t(74) = 2.74$, $p = .008$. Finally, regressing free recall performance on both experimental condition and mHR rendered mHR significant ($\beta = -.18$, 95% CI [−.31, −.06], $t(58) = -2.76$, $p = .008$) but not experimental condition, $\beta = 4.81$, 95% CI [−1.06, 10.33], $t(58) = -1.69$, $p = .09$. The indirect effect of the independent variable (role) on the dependent variable (correct free recall) via

Table 1
Means, SDs, and 95% Confidence Intervals (CIs) for Maximum Heart Rate (MHR) and Heart Rate Variability (HRV) by Experimental Group

	Active Officers		Observer Officers	
	Mean (*SD*)	95% CI	Mean (*SD*)	95% CI
Max HR (mHR)**	126 (19.95)	[119.38, 132.89]	110 (22.58)	[102.44, 118.94]
HR variability (HRV)*	22.34 (14.89)	[17.81, 27.51]	32.83 (20.63)	[25.58, 40.35]

* $p < .05$. ** $p < .01$.

Table 2
Means, SDs, and 95% Confidence Intervals (CIs) for Correct and Incorrect Items Reported and Accuracy Rate by Experimental Group

	Active Officers		Observer Officers	
	Mean (*SD*)	95% CI	Mean (*SD*)	95% CI
Briefing phase				
Correct	47.46 (17.37)	[41.97, 52.68]	47.24 (16.01)	[42.32, 52.45]
Incorrect	4.72 (5.66)	[3.16, 6.86]	4.54 (4.21)	[3.25, 6.00]
Accuracy rate	.91 (.08)	[.88, .94]	.92 (.07)	[.90, .94]
Critical response phase				
Correct**	38.67 (10.17)	[35.61, 41.97]	45.84 (12.59)	[42.02, 50.03]
Incorrect	1.64 (2.18)	[1.00, 2.40]	1.29 (1.05)	[.97, 1.66]
Accuracy	.96 (.04)	[.95, .97]	.97 (.02)	[.96, .98]

** $p < .01$.

the mediator was significant, Sobel Test = 2.08, $p = .02$. Baron and Kenny (1986) state that a variable *M* functions as a mediator when the significant effect of *X* is rendered nonsignificant after controlling for *M*. Thus, in the current analyses, officer role was related to the degree of arousal experienced during the critical response phase and the effect of role on correct information recalled was mediated by arousal.

Although regressing HRV on experimental condition was statistically significant, $\beta = 10.49$, 95% CI [1.41, 19.11], $t(59) = -2.30$, $p = .025$, HRV did not mediate the effect of role on recall. Experimental condition continued to predict free recall performance ($\beta = 7.65$, 95% CI [1.52, 14.06], $t(58) = 2.59$, $p = .012$, but HRV did not, $\beta = -.007$, 95% CI [−.17, .15], $t(58) = -0.09$, $p = .93$ (see Discussion).

Closed Questions

The overall accuracy rate for the closed questions was 57% ($SD = 11.25$) with accuracy ranging from 28% to 76% and no difference between conditions although there was a trend to greater accuracy, in terms of overall accuracy rate, in the observer condition (Active $M = .55$, $SD = .13$, 95% CI [.51, .59]; Observer $M = .60$, $SD = .08$, 95% CI [.57, .62]), $t(73) = -1.78$, $p = .08$, $d = 0.46$, 95% CI [−0.002, 0.92]. On average, officers wrote a Do Not Know response for 17% of the closed questions with no difference between experimental groups in the mean frequency of Do Not Know responding (Active $M = .19$, $SD = .11$, 95% CI [.16, .23]; Observer $M = .15$, $SD = .08$, 95% CI [.12, .18]), $t(73) = 1.62$, $p = .11$, $d = 0.42$, 95% CI [−0.05, 0.88]. With respect to Do Not Know responses, participants were asked to categorize such responses to one of four categories. Notably, there was a significant effect of Officer Role on the frequency of selection of one of the categories (Category A: "*I did not report an answer because the information was not present in the event*;" see Table 3).

Target questions. Recall that the target questions were a subset of the closed questions identified by legal and police training advisors as critical from an investigative perspective. For analysis, the average number of correct, incorrect and Do Not Know responses were calculated for each of the three subsets, Precursor Perpetrator Action, Officer Response, and Perpetrator Weapon. For Perpetrator Action and Officer Response, there were no significant differences between conditions for mean number of correct, incorrect or DK responses (see Table 4).

With respect to questions pertaining to the perpetrator's weapon, there was no difference between conditions for the number of questions answered correctly or with a Do Not Know response. However, there was a significant difference between conditions for incorrect responses (see Table 4) and accuracy rate (Active $M = .55$, $SD = .33$, 95% CI [.44, .66]; Observer $M = .71$, $SD = .31$, 95% CI [.60, .81]), $t(73) = 2.13$, $p = .04$, $d = 0.50$, 95% CI [0.03, 0.96].

Memory for Weapon-Related Details

Of the 39 participants placed in the role of the Active Officer, 33 of them discharged their weapon during the scenario (85%). We conducted an additional coding of the free recall data to examine how

Table 3
Mean Proportion of Do Not Know (DK) Responses Allocated to Each Category by Experimental Group

	Active		Observer				
DK response	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI	*t* statistic[a]	*d*	95% CI
Did not report an answer because the information was not present (A)*	.17 (.24)	[.09, .27]	.05 (.16)	[.01, .12]	2.22 ($p = .03$)	.59	[.05, 1.12]
Did not report an answer because I could not recall the specific information (B)	.55 (.32)	[.44, .67]	.57 (.39)	[.43, .70]	−.22 ($p = .83$)	.05	[−.58, .47]
Did not report an answer because I truly do not remember (C)	.18 (.21)	[.11, .26]	.28 (.34)	[.16, .42]	−1.34 ($p = .19$)	.36	[−.88, .17]
Did not report an answer because I was not confident that it was be correct (D)	.09 (.15)	[.04, .15]	.09 (.19)	[.04, .17]	−.005 ($p = .99$)	.00	[−.52, .52]

[a] $df = 56$ for a between-subjects *t*-test. CI = confidence interval.
* $p < .05$.

Table 4
Mean Number of Target Questions Answered Correctly, Incorrectly, and Using A Do Not Know Response by Participant Role and Question Category

	Active		Observer				
Target questions	*M* (*SD*)	95% CI	*M* (*SD*)	95% CI	*t* statistic[a]	*d*	95% CI
Perpetrator action							
Correct	3.94 (2.11)	[3.27, 4.58]	4.43 (1.92)	[3.75., 5.03]	−1.04 ($p = .30$)	.24	[.21, .70]
Incorrect	3.49 (2.11)	[2.84, 4.12]	2.70 (1.80)	[2.11, 3.29]	1.62 ($p = .11$)	.40	[−.5, .86]
Do not know	2.61 (2.04)	[2.00, 3.25]	2.2 (1.63)	[1.69, 2.77]	.94 ($p = .35$)	.22	[−.23, .68]
Officer response							
Correct	2.51 (1.09)	[2.16, 2.85]	2.37 (1.86)	[2.00, 2.736]	.51 ($p = .61$)	.10	[−.36, .55]
Incorrect	.72 (.82)	[.45, .97]	.70 (.70)	[.48, .91]	.09 ($p = .93$)	.02	[−.43, .48]
Do not know	.74 (.91)	[.49, 1.03]	.54 (.90)	[.28, .83]	.98 ($p = .33$)	.22	[−.23, .67]
Weapon position							
Correct	2.21 (1.30)	[1.83, 2.58]	2.73 (1.30)	[2.26, 3.13]	1.75 ($p = .08$)	.41	[.06, .86]
Incorrect[*]	1.15 (1.08)	[.82, 1.51]	.65 (.88)	[.37, .95]	2.21 ($p = .03$)	.51	[.04, .97]
Do not know	.64 (1.01)	[.36, .97]	.48 (.90)	[.23, .78]	.70 ($p = .48$)	.16	[−.29, .62]

[a] $df = 74$ for a between-subjects *t*-test. CI = confidence interval.
[*] $p < .05$.

many participants spontaneously reported that the perpetrator pointed a weapon at them/in their direction in response phase (note that the gun had remained in the perpetrators waistband throughout that scene). Overall, 18% of participants spontaneously reported that a gun was pointed at them in the final part of the scenario. There was no association between Officer Role and likelihood of reporting the perpetrator pointing a gun toward the officers in the final scenario: 15% of Active officers and 22% of Observer officers reported seeing a gun in the hands of the perpetrator, $\chi^2 < 1$.

Discussion

Operationally active witnesses did not differ from their observer counterparts with respect to their recall of the briefing phase encountered before immersion in the critical response phase. Recall of the initial briefing phase arguably represents a baseline recall measure as arousal levels would have been equivalent at time of encoding (and our baseline HR measures suggest there are no other systematic physiological differences between the groups). However, an interesting difference emerged between active and observer officers in their recall of the critical response phase. Operationally active witnesses reported significantly fewer correct details about the scenario than observer witnesses. However, there were no differences, according to role type, in the overall accuracy rate of information reported. Thus, operational witnesses, in free-recall tasks at least, were able to sustain the accuracy of their accounts. Indeed, accuracy rates for freely reported information were very high across both operational and nonoperational witnesses (≥92%). In light of the high profile political, legal, and investigative contexts such accounts are evaluated in, the adoption of a conservative reporting strategy is not particularly surprising, and would explain the high levels of accuracy observed here (see Hope et al., 2013 for further discussion of this issue). Nonetheless, that operationally active witnesses reported significantly fewer correct details about the critical response phase than their nonoperational observer counterparts is important.

Our physiological data may help account for this finding. Operationally active witnesses showed significantly higher levels of physiological arousal, as marked by higher HRs and lower HRV, during the critical response phase of the scenario in comparison with their nonoperational counterparts. It is noteworthy that significantly different HR measures were recorded for active witnesses despite the fact that officers in both active and observer roles were exposed to the same critical scenario and stood side by side while the scenario unfolded. In other words, higher HRs and lower HRV did not reflect increased *physical* activity—in fact, we deliberately limited the potential for differential physical movement through the enforcement of predetermined containment positions. Therefore, active witnesses in the current study experienced higher levels of physiological arousal or stress response as a function of the demands of their operational response role.

Mediational analyses revealed that the observed effects of role on free recall performance were related to level of stress, as indexed by the maximum HR recorded during the critical response phase. Thus, while it is important to consider the role of a witness to an event, the degree of arousal experienced is also an important factor. Although HR variability recorded during the critical response phase reflected relatively increased stress workload for active witnesses, a similar meditational relationship was not observed. This result can most likely be accounted for by the sampling period for both measurements. HR typically peaked at the critical final moments in the scenario, in the same 2–3 sec, when the perpetrator turned toward the officers (i.e., the point at which 85% of active officer discharged their weapon). As such, the maximum HR measure reflects a specific instantaneous period of elevated stress. Conversely, HRV was calculated, in line with conventions over a longer period during the response phase (1 min; see Spierer, Griffiths, & Sterland, 2009). As such we would not necessarily expect HRV to mediate recall performance across this longer period. Nonetheless, it is noteworthy that this measure reflected increased stress for the active witnesses across the critical response phase.

Although high accuracy rates were obtained for both groups in these free recall reports, requiring officers to respond to closed questions produced a very different pattern of results. Accuracy rates for closed questions were comparatively low for both Active and Observer Officers—the average accuracy rate was 57% meaning that just under half of the questions were answered incorrectly or with a Do Not Know response. More important, there was a difference in response accuracy for information pertaining to the perpetrator's weapon (i.e., a legally relevant subset of questions) with active officers significantly more likely to provide incorrect information than observer officers.

These are important findings and, to our knowledge, this is the first study to document a physiological difference between witnesses who have different roles in responding to the same incident, observe that the effect of role operates through an arousal mechanism and demonstrate differences in memory performance for operational versus nonoperational witnesses. As such, these findings confirm the merits of considering the role of "operational witness" when evaluating their statements.

Despite showing a physiological profile consistent with an increased stress response, active officers in the current study clearly did not experience a generalized "catastrophic failure" of memory (in terms of the overall quantity of information reported). However, consistent with previous findings (e.g., Hope et al., 2013) they did report significantly less correct information than nonoperational observer counterparts. Examination of the errors made by active witnesses in response to closed questioning highlights particular areas of vulnerability in their recollection of the incident. Specifically, as illustrated by performance on target questions, the recall performance of active witnesses was significantly impaired, relative to observer officers, for critical information about the weapon in the final moments of their interaction with the perpetrator (i.e., when threat level was greatest). Active witnesses reported less information than observers in response to questions about the weapon and their responses were less accurate. Active officers were also more likely than their observer counterparts to categorize their use of a Do Not Know response as "*I did not report an answer because the information was not present in the event*". This suggests that details of the event were either not encoded in the first place or were no longer accessible. At first glance, the findings of the current study appear to be inconsistent with classic "attentional narrowing" accounts that propose a recall advantage for central or important stimulus information over peripheral information (e.g., Safer, Christianson, Autry, & Österlund, 1998). However, this would be a premature and likely inaccurate conclusion—particularly in the absence of data for a control, nonarousing version of the scenario. Observer officers in the current study were exposed to the same arousing encounter and displayed elevated HRs during the critical scenario (averaging 112 bmp) contrasting with their baseline HRs (averaging 67 bpm). However, this group achieved reasonably high accuracy rates for questions pertaining to the weapon, almost certainly a high priority stimulus and focus of attention during the critical response phase. There are a few potential explanations. First, it may be that observers simply did not meet an arousal threshold likely to impair recall. It is worth noting that the average HR for observers is close to the 110 bpm threshold beyond which the sympathetic nervous system (SNS) is triggered (cf. the average rate of 126 bpm for active witnesses that is well beyond this threshold; see Woody & Szechtman, 2011). Therefore, the level of arousal experienced by observers, in the absence of further competing demands, did not necessarily impede the processing of important information in the scene and, thus, the findings are likely to be consistent with previous literature. As such, moderate arousal as a function of merely witnessing a threatening incident was not responsible for reduced recall performance of active witnesses (relative to observers) in the current study. Future research might consider innovative methodological approaches where the meaning of the important stimuli is altered depending on environmental context (e.g., Pickel, 1999) and arousal level to explore these comparisons further for operationally active and nonoperational observer witnesses.

Operationally active witnesses did not show a recall enhancement for critical stimuli—in fact, relative to observer witnesses, they reported fewer correct details about the critical response phase and had poorer accuracy for questions about the weapon. There are a number of possibilities as to why this might be the case. First, in light of the physiological results, it is possible that active witnesses experienced greater stress responses than their observer counterparts that may have contributed to memory impairment for details of this final phase of the scenario where the threat level (and likely associated stress) was highest. This pattern of results is consistent with the inverted U-shaped curve predicted by arousal theories to account for performance decrements when moderate stress levels are exceeded (Morley & Farr, 2012). Second, additional cognitive load as a function of response role may have impaired their ability to process information about details of the final scene. This deficit is particularly evident for information pertaining to the fate of the perpetrator's weapon. Given that operational demands were likely to be at their greatest at this moment (i.e., attempts to attenuate immediate lethal threat), these findings are unsurprising and likely reflect reduced information processing resources. Third, the interaction between heightened stress and additional cognitive load associated with responding and attempting to neutralize the threat, may have produced impaired encoding (and/or retention) of details of the scene. As such, our findings are consistent with Morgan and Southwick (2014) who have argued on the basis of data obtained in challenging military training contexts, that memory for stressful events can be vulnerable to error and suggestion—contrary to predictions that emotionally arousing events will be remembered better than neutral events (see also Engelhard, van den Hout, & McNally, 2008; Lommen, Engelhard, & van den Hout, 2013; Morgan et al., 2013). In the current study, the increased stress experienced by active witnesses may have led to more generalized processing and resulted in the rapid extraction of gist information (e.g., Payne et al., 2002; see also Qin, Hermans, van Marle, & Fernańdez, 2012) that they were able to report accurately in a free recall account, but that was not sufficiently specific to produce accurate detailed information about the weapon when probed by closed questions.

The current data might also be considered in light of compensatory control models (e.g., Hockey, 1997). When processing resources are compromised (e.g., because of threat), individuals make strategic adjustments in the allocation of resources to maintain performance on high priority task goals (Hockey, 1997). Such adjustments often produce decrements on secondary tasks or amplify trade-offs (Hockey, 1993; Hockey & Hamilton, 1983). It has typically been assumed, here and in other research, that details associated with a deadly weapon constitute the critical stimuli. However, there are important contextual factors to take into account. First, training guidance on firearms practice (e.g., Armed Policing Authorised Professional Practice, 2013) often recommends aiming to strike central body mass depending on the type of weapon being used and training received, which means that rather than sustaining their attention on a weapon they already know to be present, officers will refocus their attention on a target's critical mass. Second, it may well be the case that once an officer has committed to a decision to discharge their weapon, other competing priorities take precedence including monitoring the immediate environment for further risk factors, securing the safety of oneself and others and planning the next action. Previous research has documented impaired recall and recognition of a critical (threatening) target individual for officers experiencing reduced processing capacity (as a function of exertion). However, detection and recall of additional risk factors in the environment was not impaired suggesting that attentional resources may have been diverted to risk assessment activities rather than the encoding of an unarmed albeit verbally threatening target (see Hope et al., 2013). A similar explanation may apply here and is consistent with the predictions of the arousal-based competition model (ABC; Mather & Sutherland, 2011). Future research should examine cognitive processing in the aftermath of weapon discharge to elucidate the relative roles of stress, cognitive load, and competing priorities in processing an incident and the subsequent impact on memory. Furthermore, researchers (and, by extension, the legal profession and evaluators) may need to a take a more contextual perspective before assuming what constitutes "critical" stimuli in a scene. As noted by Mather and Sutherland (2011) "priority is determined by bottom-up perceptual salience and top-down relevance" (p. 19), both of which would have been determined by individuals juggling competing operational demands in the current study.

The finding that 18% of the sample (largely equivalent across both Active and Observer witnesses) reported that the perpetrator pointed a weapon at them in the final critical scene is interesting—albeit predicted by expectation-driven processing. It is also consistent Morgan et al.'s finding that soldiers made nontrivial errors for the presence of weapons under stressful conditions (e.g., 27% of soldiers falsely reported that their interrogator wielded a weapon; Morgan et al., 2013). However, there are at least two other potential explanations that might account for this error. The first concerns a relatively simple visual effect comprising two stages—a feed-forward sweep that allows the rapid extraction of features from a visual scene followed by recurrent processing to produce a conscious experience (see Lamme, 2006). It may be that the expectation of a weapon in combination with visual processing under time pressure failed to detect that a weapon was not present, leaving (some) witnesses with the belief that they had in fact seen a weapon without the opportunity to fully process the scene. Alternatively, it may be that the reporting of a weapon simply reflects a reconstructive error in light of outcome knowledge. All officers who reported seeing a weapon pointed at them also discharged their own weapon. Knowing this outcome of their own decision-making, some almost certainly felt under pressure to justify the use of lethal force in a legally consistent manner. This is an interesting distinction that is important for investigators to consider in the evaluation of accounts of firearms incidents. Generally, officers will have made a decision to discharge their weapon as a function of their contemporaneous perception and sense-making at the scene (i.e., "What I think I see/What I think is happening") whereas their account justifying their actions at the scene is likely to be based on retrospective and necessarily reconstructive processes (i.e., "What I think I must have seen/What must have happened"). Just as legal decisions in the aftermath of a shooting are made with the benefit of hindsight and biases associated with the presence of outcome information (Villejoubert, O'Keeffe, Alison, & Cole, 2006), officers' post hoc evaluations may be similarly vulnerable to the influence of hindsight and outcome information. In summary, it is not possible to determine whether this error reflects a memory distortion or a post hoc justification informed by outcome bias. Similarly, it cannot be easily attributed to attentional phenomena (e.g., "inattentional blindness"; Chabris, Weinburger, Fontaine, & Simons, 2011) in the absence of relevant measurement data—particularly given exposure duration, the error reporting rate, and distribution of the error across active and observer conditions. Nonetheless, during debriefing, officers who had reported seeing a weapon expressed surprise when told that was not possible. Therefore, further research is necessary to disentangle the cognitive and social effects producing this erroneous reporting under stressful conditions. Indeed, the current research only considered the effects of stress at encoding on output. Given the high stakes of real-life shooting investigations, it is reasonable to assume that officers may also experience stress at the reporting stage.

Although officers took part in a challenging, tightly controlled, simulated incident that replicated high-quality training scenarios, the experience was obviously not as dangerous or consequential as a real-world incident involving lethal weapons. Furthermore, we only ran participants in one stimulus event. However, little research in this particular applied context, with the exception of studies conducted by Morgan and his colleagues (2013, 2014) in a military setting, has achieved similar levels of ecological validity (see Hope et al., 2012, 2013 for further discussion of this issue). Our decision to recruit pairs of officers was deliberate to limit any extraneous effects of law enforcement training and experience on recall performance. Of course, it is possible that the performance of the officer designated observer status does not necessarily replicate the performance of a lay, bystander witness. Future research should consider the extent to which expertise or domain knowledge held by operational witnesses contributes to their subsequent recall of incident and explore whether this knowledge can be capitalized on to support retrieval for operational incidents. It should also be noted that the current research involved a single White male target—perpetrator race was not manipulated and nor was "decision to shoot" a key dependent variable. Therefore, the current research cannot speak to racial aspects of recent high profile police shootings in the United States, or indeed elsewhere (for recent research on race and shooting behavior, see Cox, Devine, Plant, & Schwartz, 2014; Sim, Correll, & Sadler, 2013). Further research, involving high fidelity simulations and methodologically rigorous experimental designs, is needed to examine the determinants of decisions to shoot in diverse policing contexts. Finally, as is common in research investigating eyewitness recall, our analytical approach necessitated a number of statistical tests, which can increase the likelihood of Type I (familywise) error. For this reason, our interpretation of our results relies heavily on measures of effect size, rather than solely the statistical significance of any finding.

Through examining the performance of witnesses who, by virtue of their duty, are required not only to witness but also to react and respond under stressful conditions, the current research constitutes an important and timely contribution both to the psychological literature and wider policy concerns in legal and investigative contexts. To date, little research has systematically examined the recall of officers for challenging or threatening operational incidents, particular those involving use of lethal force. However, the investigation of such incidents constitutes a major and high profile task both for police forces and external agencies such as the Independent Police Complaints Commissions (IPCC) established in the U.K. under the Police Reforms Act, 2002. Internationally, such investigations are typically high profile, attracting both public and media attention, and have serious consequences for the officers involved (Goodwill et al., 2010). Thus, the development of evidence-based policy and investigative practice is critical. The current results document the vulnerability of memory in this context and highlight the need for well-informed approaches to eliciting information from operationally active witnesses.

References

ACPO (Association of Chief Police Officers; 2011). *Manual of guidance on the management, command and deployment of armed officers* (3rd ed.). Retrieved from http://www.acpo.police.uk/documents/uniformed/2011/201111MCDofAO3.pdf

Adolphs, R., Tranel, D., & Buchanan, T. W. (2005). Amygdala damage impairs emotional memory for gist but not details of complex stimuli. *Nature Neuroscience, 8,* 512–518.

Akinola, M., & Mendes, W. B. (2012). Stress-induced cortisol facilitates threat-related decision making among police officers. *Behavioral Neuroscience, 126,* 167–174. http://dx.doi.org/10.1037/a0026657

Alferes, V. R., & Kenny, D. A. (2009). SPSS programs for the measurement of nonindependence in standard dyadic designs. *Behavior Research Methods, 41,* 47–54. http://dx.doi.org/10.3758/BRM.41.1.47

Alpert, G. P. (2009). Interpreting police use of force and the construction of reality. *Criminology & Public Policy, 8,* 111–115. http://dx.doi.org/10.1111/j.1745-9133.2009.00536.x

Armed Policing Authorised Professional Practice. (2013). *Armed policing*. Retrieved from http://www.app.college.police.uk/app-content/armed-policing/?s=

Banks, J. B., Tartar, J. L., & Welhaf, M. S. (2014). Where's the impairment: An examination of factors that impact sustained attention following a stressor. *Cognition and Emotion, 28,* 856–866. http://dx.doi.org/10.1080/02699931.2013.857643

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. http://dx.doi.org/10.1037/0022-3514.51.6.1173

Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, Judges and Law Enforcement to eyewitness experts. *Applied Cognitive Psychology, 20,* 115–129. http://dx.doi.org/10.1002/acp.1171

Betts, K. R., & Hinsz, V. B. (2013). Strong shared representations promote schema-consistent memory errors in groups. *Group Processes & Intergroup Relations, 16,* 734–751. http://dx.doi.org/10.1177/1368430213486206

Braidwood, T. R. (2010). *Why? The Robert Dziekanski Tragedy: Braidwood commmission on the death of Robert Dziekanski (B. C.)*. ISBN 978-0-7726-6252-1. Retrieved from http://www.braidwoodinquiry.ca/report/P2Report.php

Chabris, C. F., Weinberger, A., Fontaine, M., & Simons, D. J. (2011). You do not talk about Fight Club if you do not notice Fight Club: Inattentional blindness for a simulated real-world assault. *Iperception, 2,* 150–153. http://dx.doi.org/10.1068/i0436

Clark, R. C. (2008). *Building expertise: Cognitive methods for training and performance improvement*. San Francisco, CA: Wiley.

Conway, M. A. (2012). Ten things the law and others should know about human memory. In L. Nadel & W. Sinnott-Armstrong (Eds.), *Memory and law* (pp. 359–372). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199920754.003.0014

Cox, W. T. L., Devine, P. G., Plant, E. A., & Schwartz, L. L. (2014). Towards a comprehensive understanding of officers' shooting decisions: No simple answers to this complex problem. *Basic and Applied Social Psychology, 36,* 356–364. http://dx.doi.org/10.1080/01973533.2014.923312

Davis, D., & Loftus, E. F. (2009). Expectancies, emotion and memory reports of visual events: Applications in the legal system. In J. R. Brockmole (Ed.), *The visual world in memory* (pp. 178–214). New York, NY: Psychology Press.

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 and Human Behavior, 28,* 687–706. http://dx.doi.org/10.1007/s10979-004-0565-x

De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., & Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews, 19,* 269–301.

deq Uervain, D. J., Roozendaal, B., Nitsch, R. M., McGaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature Neuroscience, 3,* 313–314. http://dx.doi.org/10.1038/73873

Dodd, V., & Travis, A. (2014). *Mark Duggan police must stop refusing to be interviewed, says IPCC head*. Retrieved from http://www.theguardian.com/uk-news/2014/jan/14/mark-duggan-police-must-talk-says-ipcc

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66,* 183–201. http://dx.doi.org/10.1037/h0047707

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology, 87,* 876–893. http://dx.doi.org/10.1037/0022-3514.87.6.876

Engelhard, I. M., van den Hout, M. A., & McNally, R. J. (2008). Memory consistency for traumatic events in Dutch soldiers deployed to Iraq. *Memory, 16,* 3–9. http://dx.doi.org/10.1080/09658210701334022

Eyre, M., & Alison, L. (2007). To decide or not to decide: Decision making and decision avoidance in critical incidents. In D. Carson, R. Milne, F. J. Pakes, K. Shalev, & A. Shawyer (Eds.), *Applying psychology to criminal justice* (pp. 211–232). Chichester: Wiley. http://dx.doi.org/10.1002/9780470713068.ch13

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C Thomas.

Goodwill, A. M., Alison, L., Lehmann, R., Francis, A., & Eyre, M. (2010). The impact of outcome knowledge, role, and quality of information on the perceived legitimacy of lethal force decisions in counter-terrorism operations. *Behavioral Sciences & the Law, 28,* 337–350.

Healy, A. F., Kole, J. A., & Bourne, L. E., Jr. (2014). Training principles to advance expertise. *Frontiers in Psychology, 5,* 131. http://dx.doi.org/10.3389/fpsyg.2014.00131

Heaton-Armstrong, A., & Wolchover, D. (1993, April). Knocking heads together. *Counsel,* 14–15.

Heaton-Armstrong, A., & Wolchover, D. (2009, January). Conferring at the crossroads. *Counsel, 18*.

Hockey, G. R. J. (1993). Cognitive-energetical control mechanisms in the management of work demands and psychological health. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention, selection. Awareness and control: A tribute to Donald Broadbent* (pp. 328–345). Oxford: Oxford University Press.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload; a cognitive-energetical framework. *Biological Psychology, 45,* 73–93. http://dx.doi.org/10.1016/S0301-0511(96)05223-4

Hockey, G. R. J., & Hamilton, P. (1983). The cognitive patterning of stress. In G. R. J. Hockey (Ed.), *Stress and fatigue in human performance* (pp. 141–177). New York, NY: Wiley.

Hope, L., Gabbert, F., & Fraser, J. (2013). Postincident conferring by law enforcement officers: Determining the impact of team discussions on statement content, accuracy, and officer beliefs. *Law and Human Behavior, 37,* 117–127. http://dx.doi.org/10.1037/lhb0000019

Hope, L., Lewinski, W., Dixon, J., Blocksidge, D., & Gabbert, F. (2012). Witnesses in action: The effect of physical exertion on recall and recognition. *Psychological Science, 23,* 386–390. http://dx.doi.org/10.1177/0956797611431463

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.

Kleider, H. M., Pezdek, K., Goldinger, S. D., & Kirk, K. A. (2008). Schema-driven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology, 22,* 1–20. http://dx.doi.org/10.1002/acp.1361

Klein, G. A. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kruglanski, A. W., & Freund, T. (1983). The freezing and un-freezing of layinferences: Effects of impressional primacy, ethnic stereotyping and numerical anchoring. *Journal of Experimental Social Psychology, 19,* 448–468. http://dx.doi.org/10.1016/0022-1031(83)90022-7

Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *The Journal of Neuroscience, 25,* 2977–2982. http://dx.doi.org/10.1523/JNEUROSCI.5139-04.2005

Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31,* 1201–1231. http://dx.doi.org/10.1016/j.chiabu.2007.03.021

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10,* 494–501. http://dx.doi.org/10.1016/j.tics.2006.09.001

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174. http://dx.doi.org/10.2307/2529310

Leach, J., & Ansell, L. (2008). Impairment in attentional processing in a field survival environment. *Applied Cognitive Psychology, 22,* 643–652. http://dx.doi.org/10.1002/acp.1385

Leach, J., & Griffith, R. (2008). Restrictions in working memory capacity during parachuting: A possible cause of 'no pull' fatalities. *Applied Cognitive Psychology, 22,* 147–157. http://dx.doi.org/10.1002/acp.1364

Lommen, M. J. J., Engelhard, I. M., & van den Hout, M. A. (2013). Susceptibility to long-term misinformation effect outside of the laboratory. *European Journal of Psychotraumatology, 4,* 19864. http://dx.doi.org/10.3402/ejpt.v4i0.19864

Lupien, S. J., & Lepage, M. (2001). Stress, memory, and the hippocampus: Can't live with it, can't live without it. *Behavioural Brain Research, 127,* 137–158. http://dx.doi.org/10.1016/S0166-4328(01)00361-8

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science, 6,* 114–133. http://dx.doi.org/10.1177/1745691611400234

McGaugh, J. L. (2013). Making lasting memories: Remembering the significant. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 10401–10407. http://dx.doi.org/10.1073/%2Fpnas.1301209110

Meyerhoff, J. L., Norris, W., Saviolakis, G. A., Wollert, T., Burge, B., Atkins, V., & Spielberger, C. (2004). Evaluating performance of law enforcement personnel during a stressful training scenario. *Annals of the New York Academy of Sciences, 1032,* 250–253. http://dx.doi.org/10.1196/annals.1314.031

Morgan, C. A., III, Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., . . . Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry, 27,* 265–279. http://dx.doi.org/10.1016/j.ijlp.2004.03.004

Morgan, C. A., III, & Southwick, S. (2014). Perspective: I believe what I remember, but it may not be true. *Neurobiology of Learning and Memory, 112,* 101–103. http://dx.doi.org/10.1016/j.nlm.2013.12.011

Morgan, C. A., III, Southwick, S., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *International Journal of Law and Psychiatry, 36,* 11–17. http://dx.doi.org/10.1016/j.ijlp.2012.11.002

Morley, J. E., & Farr, S. A. (2012). Hormesis and amyloid-β protein: Physiology or pathology? *Journal of Alzheimer's Disease, 29,* 487–492.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review, 112,* 217–242. http://dx.doi.org/10.1037/0033-295X.112.1.217

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA: Freeman.

Nieuwenhuys, A., Savelsbergh, G. J. P., & Oudejans, R. R. (2012). Shoot or don't shoot? Why police officers are more inclined to shoot when they are anxious. *Emotion, 12,* 827–833. http://dx.doi.org/10.1037/a0025699

Payne, J. D., Jackson, E. D., Ryan, L., Hoscheidt, S., Jacobs, J. W., & Nadel, L. (2006). The impact of stress on neutral and emotional aspects of episodic memory. *Memory, 14,* 1–16. http://dx.doi.org/10.1080/09658210500139176

Payne, J. D., Nadel, L., Allen, J. J., Thomas, K. G., & Jacobs, W. J. (2002). The effects of experimentally induced stress on false recognition. *Memory, 10,* 1–6. http://dx.doi.org/10.1080/09658210143000119

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology, 57,* 27–53. http://dx.doi.org/10.1146/annurev.psych.56.091103.070234

Pickel, K. L. (1999). The influence of context on "weapon focus" effect. *Law and Human Behavior, 23,* 299–311. http://dx.doi.org/10.1023/A:1022356431375

Plant, K., & Stanton, N. A. (2012). Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. *Safety Science, 50,* 300–315. http://dx.doi.org/10.1016/j.ssci.2011.09.005

Qin, S., Hermans, E. J., van Marle, H. J. F., & Fernández, G. (2012). Understanding low reliability of memories for neutral information encoded under stress: Alterations in memory-related activation in the hippocampus and midbrain. *The Journal of Neuroscience, 32,* 4032–4041. http://dx.doi.org/10.1523/JNEUROSCI.3101-11.2012

Roozendaal, B. (2000). Glucocorticoids and the regulation of memory consolidation. *Psychoneuroendocrinology, 25,* 213–238. http://dx.doi.org/10.1016/S0306-4530(99)00058-X

Safer, M. A., Christianson, S.-Å., Autry, M. W., & Österlund, K. (1998). Tunnel memory for traumatic events. *Applied Cognitive Psychology, 12,* 99–117. http://dx.doi.org/10.1002/(SICI)1099-0720(199804)12:2<99::AID-ACP509>3.0.CO;2-7

Schacter, D. L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *American Psychologist, 54,* 182–203. http://dx.doi.org/10.1037/0003-066X.54.3.182

Schoofs, D., Wolf, O. T., & Smeets, T. (2009). Cold pressor stress impairs performance on working memory tasks requiring executive functions in healthy young men. *Behavioral Neuroscience, 123,* 1066–1075. http://dx.doi.org/10.1037/a0016980

Scoboria, A., & Fisico, S. (2013). Encouraging and clarifying "don't know" responses enhances interview quality. *Journal of Experimental Psychology: Applied, 19,* 72–82. http://dx.doi.org/10.1037/a0032067

Shackman, A. J., Sarinopoulos, I., Maxwell, J. S., Pizzagalli, D. A., Lavric, A., & Davidson, R. J. (2006). Anxiety selectively disrupts visuospatial working memory. *Emotion, 6,* 40–61. http://dx.doi.org/10.1037/1528-3542.6.1.40

Sharman, S. J., & Powell, M. B. (2012). Comparison of adult witnesses' suggestibility across various types of leading questions. *Applied Cognitive Psychology, 26,* 48–53. http://dx.doi.org/10.1002/acp.1793

Sherman, J. W., & Bessenoff, G. R. (1999). Stereotypes as source monitoring cues: On the interaction between episodic and semantic memory. *Psychological Science, 10,* 106–110. http://dx.doi.org/10.1111/1467-9280.00116

Sim, J. J., Correll, J., & Sadler, M. S. (2013). Understanding police and expert performance: When training attenuates (vs. exacerbates) stereotypic bias in the decision to shoot. *Personality and Social Psychology Bulletin, 39,* 291–304. http://dx.doi.org/10.1177/0146167212473157

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the U.S. population. *PLoS ONE, 6,* e22757. http://dx.doi.org/10.1371/journal.pone.0022757

Spierer, D. K., Griffiths, E., & Sterland, T. (2009). Fight or flight: Measuring and understanding human stress response in tactical situations. *The Tactical Edge*. Retrieved from http://www.polar.com/files/pdf/Flight_or_Fight_-_Using_Heart_Rate_and_Heart_Rate_Variability_to_Measure_Bomb_Technician_Stress.pdf

Taverniers, J., Taylor, M. K., & Smeets, T. (2013). Delayed memory effects after intense stress in Special Forces candidates: Exploring path processes between cortisol secretion and memory recall. *Stress, 16,* 311–320. http://dx.doi.org/10.3109/10253890.2012.721824

Thayer, J. F., Ahs, F., Fredrikson, M., Sollers, J. J., III, & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience and Biobehavioral Reviews, 36,* 747–756. http://dx.doi.org/10.1016/j.neubiorev.2011.11.009

Tuckey, M. R., & Brewer, N. (2003). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied, 9,* 101–118. http://dx.doi.org/10.1037/1076-898X.9.2.101

Villejoubert, G., O'Keeffe, C. J., Alison, L. J., & Cole, J. C. (2006). Hindsight bias and shooting incidents. In S. Giles & M. Santarcangelo (Eds.), *Psychological aspects of legal processes* (pp. 17–24). Liverpool: IA-IP Press.

Vrij, A., Hope, L., & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights from Behavioral and Brain Sciences, 1,* 129–136. http://dx.doi.org/10.1177/2372732214548592

Wolf, O. T. (2003). HPA axis and memory. *Best Practice & Research Clinical Endocrinology & Metabolism: Clinical and Experimental, 17,* 287–299. http://dx.doi.org/10.1016/S1521-690X(02)00101-X

Wolf, O. T., Schommer, N. C., Hellhammer, D. H., McEwen, B. S., & Kirschbaum, C. (2001). The relationship between stress induced cortisol levels and memory differs between men and women. *Psychoneuroendocrinology, 26,* 711–720. http://dx.doi.org/10.1016/S0306-4530(01)00025-7

Woody, E. Z., & Szechtman, H. (2011). Adaptation to potential threat: The evolution, neurobiology, and psychopathology of the security motivation system. *Neuroscience and Biobehavioral Reviews, 35,* 1019–1033. http://dx.doi.org/10.1016/j.neubiorev.2010.08.003

Received September 17, 2014
Revision received August 14, 2015
Accepted August 16, 2015 ■