Perceptions of officer-involved shootings by police officers versus civilians

Kathy Pezdek, Tyler Shapland & Jessica Barragan

To cite this article: Kathy Pezdek, Tyler Shapland & Jessica Barragan (22 Feb 2024): Perceptions of officer-involved shootings by police officers versus civilians, Psychology, Crime & Law, DOI: 10.1080/1068316X.2024.2319313

To link to this article: https://doi.org/10.1080/1068316X.2024.2319313

Published online: 22 Feb 2024.
Perceptions of officer-involved shootings by police officers versus civilians
Kathy Pezdek, Tyler Shapland and Jessica Barragan

Department of Psychology, Claremont Graduate University, Claremont, CA, USA

ABSTRACT
In two preregistered experiments, we explore how perception and memory for a use-of-force incident differ between officers who participated in the incident live and civilians who later viewed a Body-Worn Camera (BWC) video of the incident. In Experiment 1, responses were compared between online civilians and officers who had participated live in a shooting simulator. Responses to event memory and state of mind questions revealed numerous differences between these two groups. Experiment 2 assessed specific mechanisms underlying these effects with an additional group of officers who participated online. Our results have important implications for the application of Graham v. Connor, 1989. This US Supreme Court decision provides the superordinate legal context for determining whether the force used by an officer was justified. It is important to acknowledge that the perspective of an officer is likely to differ from that of civilians who only view the officer’s BWC recording afterward.

The situation is a common one. A police officer receives a dispatch call, goes to the specified location and is then involved in a confrontation with a civilian. At the extreme, such confrontations might involve a use of force by an officer. Nowadays, police Body-Worn Cameras (BWC) routinely record these incidents. Later, civilians observe the BWC footage of the incident, typically on television or social media. There are frequently differences between the participating officer and civilian observers in their account of what transpired. Some of these differences might stem from, for example, biases against police held by civilians (Ewanation et al., 2022) or officers’ fear of being charged with excessive use-of-force, a terminable offense. However, there are also cognitive factors that are likely to differentially affect perceptions of and memory for the participating officer and civilian observers who later only view the police BWC footage. Assessing whether the behavior of participating officers was reasonable relies on understanding these cognitive differences.

In this study, we address two issues. First, in what ways do perception of and memory for a simulated use-of-force incident differ between police officers who participated in the incident and civilians who later only viewed the BWC recording of the incident? Second, what mechanisms can account for these cognitive processing differences?
This research has important legal implications. The US Supreme Court decision of *Graham v. Connor* (1989) provides the superordinate legal context in the United States, for determining whether the force used by an officer was justified. In this ruling, the court concluded that a test criterion of ‘reasonableness’ must be used when assessing the actions of law enforcement officers in use-of-force incidents. This ‘reasonableness’ criterion for law enforcement was defined as being ‘judged from the perspective of a reasonable officer on the scene, rather than with the 20/20 vision of hindsight’ (*Graham v. Connor*, 1989, p. 396). Importantly, if the perspective of an officer on the scene differs from that of civilians (including triers of fact) who later only view the officer’s BWC footage, then the civilians are not likely to be using the same standard used by the officer, thus distorting the application of the principal legal ruling in *Graham v. Connor* (1989). Whereas participating officers are more likely to judge a use-of-force incident using the ‘reasonableness criterion’, in contrast, a significant portion of American civilians express disapproval of legally reasonable police use-of-force, and increasingly so over the past several decades (Celestin & Kruschke, 2019; Mortgos & Adams, 2020). There are several cognitive reasons for predicting differences between officers and civilians in their perceptions of use-of-force. Here we consider four proposed mechanisms for this effect.

The first proposed cognitive mechanism considered is the degree of schematic processing related to information in the dispatch call. Officers are typically called to a critical incident with a dispatch call that succinctly conveys what is known about the situation and people involved. This information shapes the officer’s expectations about the level of danger likely to be encountered. The Police Executive Research Forum [PERF] (2016) provides a decision-making model to guide officers’ response to a critical incident. The first two of the five steps in this model involve (1) collecting information, and (2) assessing the situation’s threat and risk. With the information in a dispatch call, the officer can make these early assessments – correctly or incorrectly – before arriving at the scene. In fact, in a randomized controlled experiment with police officers in a firearms simulator, Taylor (2019) reported that when deciding whether to pull the trigger, police rely heavily on dispatch information. Without the information in the dispatch call, civilians are likely to view the BWC footage with different expectations than those of the officer, and these expectations will shape their cognitive processing of the incident.

Cognitive psychologists have long known that people’s expectations, schemata and frameworks shape how they encode information, which in turn affects comprehension and memory (Rumelhart & Ortony, 1977; Schank & Abelson, 1977). Further, similar to Bartlett’s (1932) classic findings regarding the effects of schemata on memory, Holst and Pezdek (1992) reported that in a mock trial setting, people erroneously ‘remembered’ script consistent details of a robbery narrative even if these details had not been presented. Similar results were reported by Kleider et al. (2008) with schema expectations manipulated by pre-existing gender stereotypes.

A second proposed mechanism to account for the effects in this study is potential differences because of the medium; that is, what is visible in a live event – even a simulated one as in the shooting simulator in our study – as compared to a BWC recording of it. Differences between an officer’s perspective and that of the BWC on his/her chest could result from differences in, for example, where the officer’s head versus the BWC were directed, the greater clarity of peripheral information for the BWC than for the
officer’s vision, and the officer’s ability to see off-camera details. Also, because of its vividness, people tend to overbelieve video evidence (in courtrooms, for example) without discriminating well between accurate and inaccurate interpretations of the information presented (Granot et al., 2018). Further, and more specifically, Turner et al. (2019) reported that lower judgments of intentionality were reported in body cam videos than dash cam videos of an incident because the focal actor was less visually salient in the former than the latter.

Third, the perspective of an officer is also likely to differ from that of a civilian viewing an officer’s BWC footage afterward because of the higher level of stress experienced by the participating officer. An extensive research literature has supported the detrimental effect of stress on memory (Wolf, 2009). Within the field of eyewitness memory, a meta-analytic review by Deffenbacher et al. (2004) reported that high stress consistently impairs the accuracy of memory. This is consistent with models by Eysenck and Calvo (1992) and others proposing that high stress restricts working memory capacity. The high stress experienced by an officer participating in a lethal force situation is also likely to be associated with perceptual distortion. This has been reported in controlled studies with officers tasked with responding to (a) a live simulated scenario involving a weapon (Hope et al., 2016), (b) an active shooter scenario (Alpert et al., 2012) and (c) a virtual reality simulation with a ‘shoot/no-shoot’ scenario (Stanny & Johnson, 2000).5

On a related point, it might be argued that memory for an event is more accurate when people actively participate in the event than when they simply witness a video of it. However, Hope et al. (2016) reported that officers who participated in a scenario reported fewer correct details about the critical phase of the scenario than those who simply observed it. Similar results with measures of memory accuracy were reported by Ihlebæk et al. (2003) and Kassin (1984).

A fourth proposed mechanism is cohort differences between officers and civilians based on the expectations and perspective derived from officers’ training and experience. This relates to the fact that police departments have policies, procedures, and guidelines for responding to use-of-force incidents, and police officers have extensive training and experience following these (see for example, Police Executive Research Forum [PERF], 2016, and therein, the PERF Critical Decision-Making Model). For example, in assessing and reassessing the best course of action in an incident, officers are typically instructed to consider the proportionality of their response, and they are specifically trained on de-escalation strategies especially communication. On the other hand, the public is rarely familiar with these policies, procedures and guidelines. Thus, while police officers are trained to have a decision-making mental model guiding their response to a critical incident, civilians view BWC footage without this framework. Critically, this difference is likely to produce cognitive differences between police officers and civilians in how they encode and remember what transpires in a use-of-force incident.

Despite the frequent contrasts made between perceptions of police officers and civilians – and the popular press and social media interest in this phenomenon – prior to this study, no research has specifically assessed whether the perceptions of and memory for a simulated use-of-force incident differ between police officers who participated in the incident and civilians who only viewed the officer’s BWC footage afterward.6 If such differences do exist, this would critically distort the application of the principal legal ruling in Graham v. Connor (1989). Assessing this is the primary purpose of this study.

5 Deffenbacher et al. (2004) reported that high stress consistently impairs the accuracy of memory. This is consistent with models by Eysenck and Calvo (1992) and others proposing that high stress restricts working memory capacity. The high stress experienced by an officer participating in a lethal force situation is also likely to be associated with perceptual distortion.

6 Ihlebæk et al. (2003) and Kassin (1984) reported similar results with measures of memory accuracy.
Experiment 1

Method

This study includes two experiments. In Experiment 1, responses from 138 civilians recruited from MTurk were compared to responses from 69 police officers from a previous study by Pezdek et al. (2022). In the previous study, experienced police officers at a police training facility participated in an Officer Involved Shooting (OIS) training simulator. They each participated in the same two training scenarios with their BWC activated and then answered questions afterward. In Experiment 1 of the current study, each civilian was randomly paired with one of these 69 officers and viewed that officer’s BWC footage from their participation in both scenarios in the previous study. Half of the civilian participants were presented the dispatch call information before viewing the BWC recording of each scenario, and half were not.

Participants and design

Experiment 1 was a three-groups design that included (a) officers with dispatch call information provided who participated live in a training simulator (N = 69), (b) civilians with dispatch call information provided who participated online (N = 69), and (c) civilians with no dispatch call information provided who participated online (N = 69). Civilian participants were randomly but equally assigned to the dispatch or no-dispatch conditions. Following the methods proposed by Faul et al. (2007), we used G*Power to conduct an a priori power analysis for the main effect of dispatch condition in an independent samples t-test with the following parameters: Power = .85, α = .05, a medium effect size9 (Cohen’s d) = .50. The officers were selected from those who participated in Experiments 1 or 2 in the previous study by Pezdek et al. (2022). All officers were from the same police department in the metropolitan Los Angeles area and had an average of 15.28 (SD = 10.52) years of service. A total of 125 officers each participated in Experiment 1 or 2 in the previous study. From these, we selected for this study, the 69 officers whose BWC footage captured each scenario most clearly, based on the judgment of two raters. In other words, we selected the officers for whom their BWC was directed toward the projection screen throughout most of the scenario and was not obscured by the participating officer’s body or arms. In our previous study, officers answered the same questions about each scenario at three points in time. Only the responses from the Time 1 test immediately after participating in the scenarios were included in the present study for comparisons with civilians who also answered the same questions immediately after viewing the scenarios. The inclusion criteria for civilians were that they be at least 18 years old and currently located in the USA. All exclusion criteria are provided in the supplementary materials. The demographic variables that describe all samples for in this study are presented in Table 1.

Procedure

Each civilian was randomly paired with one of the 69 officers and viewed that one officer’s BWC footage from their participation in both scenarios in the previous study. This procedure was used to increase the range and diversity of the BWC videos included in the study. Before viewing the BWC recording of each of the two scenarios, half of the civilian
participants were presented the dispatch call information that had been presented to all officers in the previous study; half were presented no dispatch call information.

**Police officers.** The officers in this study, had participated one at a time in the training simulator that was also used in the previous study by Pezdek et al. (2022). The simulator used was Ti Training’s RECON Core. Details about this simulator system are available with the supplementary materials available online. Although the scenarios took place on a two-dimensional screen, the simulator had a shot detection system that allowed officers to move freely around the training room and still have their weapons registered by the system.

Once an officer entered the simulation room, they were given a brief description of what would transpire. The researcher emphasized that throughout, the officer needed to speak and behave as if each dispatch call was real, including using verbal de-escalation. Next, they were fitted with a vest on which a BWC was mounted. The researcher then directed the officer to their simulation equipment: gun, taser, pepper spray, and safety glasses, and instructed them to place the equipment on their person as they would with their real equipment.

Officers were then instructed to stand at the starting point, a blue arrow on the floor, but were told that they could freely move around the room once the simulation began. To increase the level of stress and encourage active engagement in the simulation, the officers were told that a shoot-back system was activated. The shoot-back system is an airsoft gun that shoots BB gun bullets controlled by the simulator to imitate a shooting gun. Although officers were told that the shoot-back system may go off during their participation, in fact, it never did. Throughout the simulation, the researcher stayed in the back of the simulation room behind a plexiglass screen in the control area.

Each officer participated in two scenarios in the simulator, counterbalanced for order of presentation. They were told that they would be responding to each scenario alone, without a partner. After participating in the first scenario, the second scenario followed immediately. After the second scenario, the officer moved to another room and was immediately presented the test questions for both scenarios. Test questions were
presented in a Qualtrics survey format on a computer, the same format on which the civilians responded.\textsuperscript{11}

\textbf{Civilian participants.} The civilians were recruited on MTurk, and each participated on their computer. They were paid $2.50 for their participation. Each of the 138 civilian participants was presented with the two BWC recordings of one of the 69 officers, again, counterbalanced for order of presentation. The same 69 BWC recordings were viewed by the 69 civilian participants in the condition in which the dispatch call information was provided and the 69 civilian participants in the no dispatch condition. Again, to clarify, each civilian participant viewed the BWC recording of both scenarios by the same officer and viewed both of these videos in the same condition, that is, either with or without the dispatch call provided (manipulated between subjects).

The civilian participants on MTurk were first read the following instructions:

Earlier this year, as part of their regular police department training, several hundred police officers participated in a shooting simulation that involved two real use of force incidents. Their participation in both use of force simulations was recorded on their body worn camera. They were told to respond to each incident as if it was a real-life threat. Today, you will view these two real body worn camera recordings of one of the officers and answer some questions about each incident afterward. Just to clarify, what you will see is a replay of the body worn camera recorded footage of the officer for each incident.

The civilians then viewed the BWC recording of both scenarios. Immediately afterward, they were presented the same sequence of 11 questions for each scenario and responded on their computer. Each scenario played out to the end and all officers shot their gun by the end of each scenario.

\textbf{Materials}

\textit{Description of scenarios.} Each officer participated in the same two scenarios in the simulator. Each scenario involved a domestic violence incident that officers reported to be startling and highly stressful. Calls of domestic violence are known to be among the most dangerous and potentially violent encountered by officers.\textsuperscript{12} Each scenario lasted about one minute. The scenarios were professionally prepared live-action videos that were made to accompany the shooting simulator. Each scenario was projected life size on a screen in the simulation room. Both actors in each scenario were White. Across officers, the assignment of each scenario to an experimental condition was counterbalanced and arranged so that each scenario was approximately equally often ordered first or second.

In the \textit{domestic dispute in the home scenario}, the dispatch call described a residential area where a couple reportedly had a physical altercation earlier, but it has now deescalated to just a verbal argument. The simulation began inside the residence; a frightened woman emerges who informs the officer that the male just got out of jail, has been drinking and has a gun. Shortly after, the inebriated male walks into the room, holding a wine bottle. He begins waving the bottle around, apologizing and saying, ‘everything is okay’. The male then leaves the room and is no longer visible for a few seconds. He reappears with one hand behind his back and begins yelling, ‘I am not going back to jail, man!’ as he pulls out a gun and attempts to shoot the officer. At this point, the officer could either (a)
shoot the male, or (b) be shot by the male. The scenario then ended. All officers in this study shot the male before he shot them.

In the domestic dispute in the parking lot scenario, the dispatch call described the officer sitting in a parking lot of a grocery store when a citizen informs them of a couple in a heated argument that may become physical. The simulation began with a couple arguing about the male being late to pick up their child. The argument continues for a while, with the couple ignoring the officer. When the male notices the officer, he walks towards him and shouts, ‘this is none of your business!’ While the male is shouting, the female reaches into her car and emerges holding a swaddled infant. The male then walks back to the female, and the argument continues. The male then reaches out for the child, and the woman pulls the baby away. They both struggle over the baby for a few seconds until the female, holding the baby, reaches behind her back and pulls out a gun. The male immediately retreats with his hands in the air while the female continues to brandish the gun for a few seconds. The female shoots the male and then immediately turns and points the gun at the officer. At this point, the officer could either (a) shoot the female, or (b) be shot by the female. The scenario then ended. All officers in this study shot the female before she shot them.

**Test questions.** Participants answered 11 questions for each of the two scenarios. These 11 questions are presented in Table 2. Responses represent retrospective reports because the officers were instructed to respond to each question from their point of view when they had actually participated in each scenario in the training simulator. The civilian participants were instructed to respond to each question from the officer’s point of view when they viewed an officer’s BWC recording of their participation in each scenario in the training simulator. The questions included (a) 5 event memory questions that probed what occurred in the BWC recording of the corresponding officer’s participation in each scenario in the training simulator and (b) 6 questions about the perceived state of mind of the officer during the incidents; 2 of these questions were open-ended recall questions, and 4 were Likert scale questions. Although we specify questions as basically tests of event memory or memory for state of mind, an officer’s state of mind is likely to be derived from their perception and memory for the event itself and not independent of it.

**Results and discussion**

The analysis plan for both experiments in this study followed that utilized in the previous study by Pezdek et al. (2022). We specifically followed the same analysis plan in both this and the previous study to maximize the opportunity for comparisons across studies. Although we are aware of criticisms of using Multivariate Analysis of Variance (MANOVA) (see for example, Huang, 2020), in fact, in this study it actually is the combination of the dependent variables that is a more meaningful outcome than the individual dependent variables on their own. This then justifies our use of MANOVA. The results were first analyzed with a One-Way MANOVA, comparing overall performance across the 3 groups on the first 7 questions that included the 5 event memory questions and the 2 open-ended state of mind questions. The MANOVA revealed that across the first 7 questions, responses significantly differed among the three groups, $F(14, 662) = 4.16$, $p < .001$; Pillai’s Trace = .162, $\eta_p^2 = .08$. 
Separate One-Way Analyses of Variance (ANOVA) were then performed on each of the 11 items separately, comparing responses among the three groups. These 11 ANOVAs allowed us to assess which measures were relatively more sensitive to between group differences in this study. Each significant ANOVA was then followed with 3 independent groups pair-wise t-tests using a Holm–Bonferroni correction to control for family-wise error rate (Aickin & Gensler, 1996). The mean response to each question in each condition is presented in Table 2. Each participant viewed and responded to two scenarios, and the results presented are averaged across the two scenarios for each participant. The statistical results for all significant and nonsignificant effects for all questions are included with the supplementary materials on the OSF site for this experiment.14

Table 2. Mean [and 95% confidence interval] for each test question in each participant group for experiment 1.

<table>
<thead>
<tr>
<th>Question</th>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. At what point did (you/the officer) first draw (your/their) gun?</td>
<td>40.35 [37.04, 43.66]</td>
<td>41.64 [39.49, 44.79]</td>
<td>41.89 [38.73, 45.05]</td>
<td>.779</td>
<td>.00</td>
</tr>
<tr>
<td>2. At what point did (you/the officer) first shoot (your/their) gun?</td>
<td>53.68 [51.29, 56.08]</td>
<td>53.22 [50.81, 55.62]</td>
<td>52.50 [50.09, 54.90]</td>
<td>.786</td>
<td>.00</td>
</tr>
<tr>
<td>3. How many times did (you/the officer) shoot (your/their) gun?</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>&lt;.001</td>
<td>.03</td>
</tr>
<tr>
<td>4. Which equipment did (you/the officer) pull out first?</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>&lt;.01</td>
<td>.03</td>
</tr>
<tr>
<td>5. What was the total duration of this event?</td>
<td>110.22 [99.59, 120.85]</td>
<td>114.88 [104.25, 125.51]</td>
<td>108.97 [98.34, 119.60]</td>
<td>.718</td>
<td>.00</td>
</tr>
</tbody>
</table>

**Event Memory Questions**

6. Assuming that this simulation was real, in one sentence specify the point at which you started to feel that (you/the officer) were in danger in this simulation.
<table>
<thead>
<tr>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.89 [29.04, 34.73]</td>
<td>38.15 [35.35, 40.96]</td>
<td>38.57 [34.74, 41.40]</td>
<td>&lt;.01</td>
<td>.03</td>
</tr>
</tbody>
</table>

7. Assuming that this simulation was real, in one sentence specify the point at which you felt that (you/the officer) were in the most danger in this simulation.
<table>
<thead>
<tr>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>45.38 [42.86, 47.90]</td>
<td>49.43 [46.96, 51.91]</td>
<td>48.50 [45.95, 51.05]</td>
<td>.064</td>
<td>.01</td>
</tr>
</tbody>
</table>

**State of Mind Questions**

8. Rate how much danger you felt (you/the officer) were in at the point that (you/you) felt the most danger in this simulation.
<table>
<thead>
<tr>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
</table>

9. Rate how much control of the situation (you/the officer) felt (you/they) had at the point that (you/they) felt the most danger in this simulation.
<table>
<thead>
<tr>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.48 [2.19, 2.77]</td>
<td>2.96 [2.68, 3.25]</td>
<td>3.10 [2.81, 3.39]</td>
<td>&lt;.01</td>
<td>.02</td>
</tr>
</tbody>
</table>

10. Now consider danger to other people in the simulation. Assuming that this simulation was real, rate how much danger you felt that other people in the were in at the point (described above) that you perceived the most danger in this simulation.
<table>
<thead>
<tr>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
</table>

11. Knowing that this was a use-of-force training session in which officers are likely to make some mistakes, rate how justified you think (you/the officer) were in the level of force that (you/the officer) chose to use in this simulation.
<table>
<thead>
<tr>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Civilians/No Dispatch</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.94 [6.71, 7.17]</td>
<td>5.83 [5.60, 6.06]</td>
<td>5.59 [5.36, 5.82]</td>
<td>&lt;.001</td>
<td>.16</td>
</tr>
</tbody>
</table>

Note. * p < .05. ** p < .01. *** p < .001. Significance tests are based on 2-tailed, one-way ANOVAs with post-hoc independent groups pair-wise t-tests using Holm–Bonferroni correction. The values for p and $\eta^2_p$ in the right-hand columns are for the one-way ANOVA. For questions 8 – 11, responses were rated on a 1 (lowest level) – 7 (highest level) Likert scale.
**Event memory questions**

For questions probing when a specific event occurred in the video, all responses were converted to time in seconds from the beginning of the video. Regarding question 1 for example, a typical officer’s response to this question in the domestic dispute in the parking lot scenario was, ‘I first drew my gun when the woman pulled out her gun and pointed it at the man’. The point in the video where the woman first pulled out her gun (and the point in the video for each specific action probed) was time stamped by two coders and this time in seconds was used as the participant’s response. Participants did not provide time estimates themselves.

(1) **At what point did (you/the officer) first draw (your/their) gun?**

The One-Way ANOVA performed on the mean response to this question in seconds yielded a nonsignificant main effect of group, $F(2, 371) = 0.25$.

(2) **At what point did (you/the officer) first shoot (your/their) gun?**

The One-Way ANOVA performed on the mean response accuracy to this question yielded a nonsignificant main effect of group, $F(2, 385) = 0.24$.

(3) **How many times did (you/the officer) shoot (your/their) gun?**

The mean number of shots fired by officers in the simulator was 3.66 [3.30, 4.01]. The One-Way ANOVA performed on the estimated mean number of shots fired yielded a significant main effect of group, $F(2, 411) = 7.12, p < .001, \eta^2_p = .03$. In their estimates of the number of shots fired, the officers differed significantly from both civilians with dispatch call, $t(411) = 3.72, p < .001, d = .45$, and civilians with no dispatch call, $t(411) = 2.40, p = .034, d = 0.29$. As can be seen in Table 2, officers were most accurate estimating the number of shots they had fired; both groups of civilians were less accurate than officers, under-estimating the number of shots that had been fired in the BWC videos they had viewed.

(4) **Which equipment did (you/the officer) pull out first?**

The choice of equipment included the gun, taser, and pepper spray. The large majority of the time the officers pulled out their gun first (73% of officers); the remainder pulled out their taser first. All officers pulled out their gun before the end of each scenario. The coding of these responses was binary, accurate or not accurate. The One-Way ANOVA performed on the mean response accuracy to this question yielded a significant main effect of group, $F(2, 411) = 7.00, p = .001, \eta^2_p = .03$. As can be seen in Table 2, officers were most accurate recalling which equipment they had pulled out first. Officers were significantly more accurate than both civilians with dispatch call, $t(411) = 3.38, p = .002, d = 0.41$, and civilians with no dispatch call, $t(411) = 3.08, p = .004, d = .37$, in recalling which equipment they had drawn first. This may simply be because the weapon drawn was not as salient in the BWC video viewed by the civilians as it was to officers participating in the actual simulator.
What was the total duration of this event?

The total duration of each event was approximately 60 s, with some variability depending on the action taken by the officer at the end of each scenario. As can be seen in Table 2, participants in all three groups over-estimated the duration of the scenarios. The One-Way ANOVA performed on the mean response accuracy to this question yielded no significant difference among the groups, $F(2, 411) = .33$.

State of mind questions

Open-ended state of mind questions. This set of analyses assessed responses to the two open-ended recall questions about the officers’ state of mind when they participated in the training simulator. For both questions, participants provided a response in terms of what event transpired in the video at the point indicated (for example, ‘I started to feel that the officer was in danger when the drunk man came into the room’). Two coders determined how many seconds from the beginning of the scenario this event occurred; each response was then converted to a time estimate in seconds. Participants did not provide time estimates themselves. On neither of these two questions did the mean ratings of the two civilian groups significantly differ.

(6) Assuming that this simulation was real, in one sentence specify the point at which you started to feel that (you/the officer) were in danger in this simulation.

The One-Way ANOVA performed on these data yielded a significant main effect of group, $F(2, 398) = 6.74, p < .001, \eta_p^2 = .03$. As can be seen in Table 2, the officers perceived that they started to feel danger significantly sooner than did the civilians. The officers started to feel danger significantly sooner than indicated by both the civilians with dispatch call, $t(398) = 3.09, p = .004, d = 0.38$, and the civilians with no dispatch call, $t(398) = 3.27, p = .003, d = 0.40$.

(7) Assuming that this simulation was real, in one sentence specify the point at which you felt that (you/the officer) were in the most danger in this simulation.

The One-Way ANOVA performed on these data yielded a nonsignificant difference among groups, $F(2, 392) = 2.77, p = .064$. Interestingly though, although officers indicated that they perceived the start of danger to be earlier in the scenarios than did civilians (responses to question 6 above), officers indicated in their responses to question 7 that they perceived peak danger at about the same time in the scenarios as civilians did; officers detected cues for start of danger earlier in the incident than did civilians but detected peak danger at the same point in time.

Likert scale state of mind questions. Four Likert scale questions assessed participants’ state of mind at various points throughout each simulation scenario. Each question was responded to on a 1 (lowest level) – 7 (highest level) Likert scale. As with all questions, the officer was responding from their own point of view when they were participating in each scenario in the simulator; civilians were asked to respond from the point of view of the participating officer in the BWC recordings they were assigned to view.
Responses to the 4 Likert scale questions about the officers’ state of mind were first analyzed with a One-Way Multivariate Analysis of Variance (MANOVA), comparing overall performance across the 3 groups. The MANOVA revealed that across these 4 questions, responses significantly differed among the three groups, $F(8, 818) = 12.10, p < .001$; Pillai’s Trace = .21, $\eta^2_p = .11$.

(8) Rate how much danger you felt (you/the officer) were in at the point that (you/they) felt the most danger in this simulation.

As presented in Table 2, participants in all three groups responded with a high rate of perceived danger as would be predicted in real-world use-of-force incidents especially those involving domestic violence. This suggests that both officers and civilians in this study found the simulations immersive and highly engaged. The One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 8.05, p < .001, \eta^2_p = .04$. The officers perceived that the incidents were significantly more dangerous than did the civilians with no dispatch call, $t(411) = 3.89, p < .001, d = 0.47$. However, the officers and the civilians with the dispatch call perceived similar levels of danger, $t(411) = 1.07, p = .284$. Interestingly, civilians who had received the dispatch call prior to viewing each video perceived a higher level of peak danger than did civilians who had received no dispatch call, $t(411) = 2.81, p = .008, d = 0.34$. Information in the dispatch call had apparently alerted civilians to a higher level of danger in each incident. This is the only measure in this study on which the two civilian groups significantly differed.

(9) Rate how much control of the situation (you/the officer) felt (you/they) had at the point that (you/they) felt the most danger in this simulation.

As can be seen in Table 2, participants in all three groups responded with a low rate of perceived control, as would be predicted in a real-world use-of-force incident. Again, this suggests that both the officers and the civilians in this study found the simulations immersive and highly engaged. The One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 5.00, p = .007, \eta^2_p = .02$. Officers perceived significantly less control in the incidents than both civilians with the dispatch call, $t(411) = 2.35, p = .039, d = 0.28$, and the civilians with no dispatch call, $t(411) = 3.01, p = .008, d = 0.36$.

(10) Now consider danger to other people in the simulation. Assuming that this simulation was real, rate how much danger you felt that other people in the video were in at the point (described above) that you perceived the most danger in this simulation.

As can be seen in Table 2, participants in all three groups responded with a high rate of perceived danger to other people in the videos, again indicating that the officers and the civilians in this study found the simulations immersive and highly engaging. The One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 9.49, p < .001, \eta^2_p = .04$. The officers perceived significantly more danger to other people than did both the civilians with the dispatch call,
Knowing that this was a use-of-force training session in which officers are likely to make some mistakes, rate how justified you think (you/the officer) were in the level of force that (you/the officer) chose to use in this simulation.

As can be seen in Table 2, participants in all three groups responded with a high rate of perceived justification for the level of force used. However, the level of justification reported by officers revealed a ceiling effect ($M = 6.94$); the officers almost always reported that they were maximally justified (i.e. provided the highest rating of 7) in the level of force they had used. Nonetheless, the One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 37.70$, $p < .001$, $\eta^2_p = .16$. Officers perceived higher levels of justification for the level of force used than did both the civilians with the dispatch call, $t(411) = 6.69$, $p < .001$, $d = 0.81$ and the civilians with no dispatch call, $t(411) = 8.14$, $p < .001$, $d = 0.98$.

**Discussion**

The major finding in this study is that officers who participated in the use-of-force training incidents and civilians who only viewed officers’ BWC recordings afterward, had quite different perceptions of and memory for what transpired. For 7 of the 11 test questions, there was a significant difference in the responses of officers and civilians. Compared to civilians, officers more accurately remembered the number of shots fired and were more accurate remembering which equipment they first pulled out. Officers rated the incidents as more dangerous to themselves and to other people, indicated that they had less control during the incident, and perceived the start of danger earlier in the incident. Officers also indicated that they were more justified in the level of force utilized.

Four mechanisms were proposed to account for differences between perceptions of and memory for a simulated use-of-force incident between police officers who participated in the incident and civilians who only viewed the officer’s BWC footage afterward. These mechanisms are: (a) degree of schematic processing related to information in the dispatch call, (b) differences in presentation medium between processing a live event versus a BWC recording of it, (c) higher stress at encoding producing more memory impairments and perceptual distortions by officers, and (d) cohort differences between officers and civilians based on expectations and perspective derived from officers’ training and experience. Based on results of Experiment 1, the role of the dispatch call can be rejected as a mechanism underlying the obtained group differences in results. Only on one of the 11 question, question 8, ratings of the peak danger, did responses differ between civilians with the dispatch call and those without the dispatch call.

In Experiment 1, comparisons of officers and civilians with dispatch call were confounded by the fact that the officers participated live in the shooting simulation and the civilians participated in an online version of this study in which they only viewed the BWC recording of an officer’s live participation. Thus, it is not possible to determine whether the differences in responses between the officers and the civilians with dispatch call were because of the cohort difference or the difference between presentation
medium, live versus online. To isolate this effect, we recruited a separate group of police officers to participate in the study online, following the procedure used by the civilians with dispatch call in Experiment 1. Experiment 2 includes this new group of officers who participated online with dispatch call, and two comparison groups from Experiment 1, police officers participating live with dispatch call and civilians participating online with dispatch call. Because the manipulation of dispatch call produced few effects in Experiment 1, in Experiment 2 each of the three groups included were presented with the dispatch call.

**Experiment 2**

**Method**

Experiment 2 was a three-groups design that included (a) responses from a new group of officers with dispatch call information provided who participated online ($N = 69$), compared to responses from two groups in Experiment 1, (b) officers with dispatch call information provided who had participated live in the shooting simulator ($N = 69$), and (c) civilians with dispatch call information provided who had participated online ($N = 69$). The new officers in Experiment 2 were recruited by sampling police chiefs from California and several other US states who were asked to distribute the Qualtrics link for our study to officers in their departments. The demographics for all four samples in this study are presented in Table 1. All procedural details were the same in Experiments 1 and 2. Following the methods proposed by Faul et al. (2007), we used G*Power to conduct an a priori power analysis for the main effect of dispatch condition in an independent samples $t$-test with the following parameters: Power = .85, $\alpha = .05$, a medium effect size ($\text{Cohen’s } d = .50$).

**Results and discussion**

In the analyses of Experiment 2, comparisons of responses between the officers who participated online and the civilians (all of whom participated online) provides an assessment of the effect of cohort (officers versus civilians). In the analyses of Experiment 2, comparisons of responses between the officers who participated live and those who participated online provides an assessment of the effect of medium (live versus online participation). Statistical comparisons between officers who participated live and civilians with dispatch call will not be reported here because this contrast was reported for these same conditions in Experiment 1, where, as discussed above, cohort and medium were confounded.

The results were first analyzed with a One-Way Multivariate Analysis of Variance (MANOVA), comparing overall performance across the 3 groups on the first 7 questions that included the 5 event memory questions and the 2 open-ended state of mind questions. The MANOVA revealed that across the first 7 questions, responses significantly differed among the three groups, $F(14, 688) = 7.16, p < .001$; Pillai’s Trace $= .254$, $\eta^2_p = .13$.

Separate One-Way Analyses of Variance (ANOVA) were then performed on each of the 11 items separately, comparing responses among the three groups. Each significant ANOVA was then followed with 2 independent groups pair-wise $t$-tests using a
Holm–Bonferroni correction to control for family-wise error rate (Aickin & Gensler, 1996). The mean response to each question in each condition is presented in Table 3. Each participant viewed and responded to two scenarios, and the results presented are averaged across the two scenarios for each participant. The statistical results for all significant and nonsignificant effects for all questions are included with the supplementary materials on the OSF site for this experiment.21

**Event memory questions**
The same method for coding event memory questions was used in Experiments 1 and 2.

Table 3. Mean [and 95% confidence interval] for each test question in each participant group for experiment 2.

<table>
<thead>
<tr>
<th>Question</th>
<th>Police Officers</th>
<th>Civilians/Dispatch</th>
<th>Online Officers</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Event Memory Questions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. At what point did (you/the officer) first draw (your/their) gun?</td>
<td>40.35 [37.04, 43.66]</td>
<td>41.64 [39.49, 44.79]</td>
<td>41.30 [38.16, 44.44]</td>
<td>.853</td>
<td>.00</td>
</tr>
<tr>
<td>2. At what point did (you/the officer) first shoot (your/their) gun?</td>
<td>53.68 [51.29, 56.08]</td>
<td>53.22 [50.81, 55.62]</td>
<td>54.04 [51.72, 56.37]</td>
<td>.887</td>
<td>.00</td>
</tr>
<tr>
<td>3. How many times did (you/the officer) shoot (your/their) gun?</td>
<td>***</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Which equipment did (you/the officer) pull out first?</td>
<td>**</td>
<td>***</td>
<td></td>
<td>&lt;.001</td>
<td>.04</td>
</tr>
<tr>
<td>5. What was the total duration of this event?</td>
<td>110.22 [99.59, 120.85]</td>
<td>114.88 [104.25, 125.51]</td>
<td>78.54 [76.63, 89.45]</td>
<td>&lt;.001</td>
<td>.06</td>
</tr>
<tr>
<td><strong>State of Mind Questions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Assuming that this simulation was real, in one sentence specify the point at which you started to feel that (you/the officer) were in danger in this simulation.</td>
<td>31.89 [29.04, 34.73]</td>
<td>38.15 [35.35, 40.96]</td>
<td>29.69 [26.84, 32.54]</td>
<td>&lt;.001</td>
<td>.04</td>
</tr>
<tr>
<td>7. Assuming that this simulation was real, in one sentence specify the point at which you felt that (you/the officer) were in the most danger in this simulation.</td>
<td>45.38 [42.86, 47.90]</td>
<td>49.43 [46.95, 51.91]</td>
<td>47.67 [45.29, 50.05]</td>
<td>.065</td>
<td>.01</td>
</tr>
<tr>
<td>8. Rate how much danger you felt (you/the officer) were in at the point that (you/the) felt the most danger in this simulation.</td>
<td>6.28 [6.06, 6.51]</td>
<td>6.11 [5.88, 6.33]</td>
<td>6.46 [6.26, 6.65]</td>
<td>&lt;.05</td>
<td>.01</td>
</tr>
<tr>
<td>9. Rate how much control of the situation (you/the officer) felt (you/the) had at the point that (you/the) felt the most danger in this simulation.</td>
<td>2.48 [2.19, 2.77]</td>
<td>2.96 [2.68, 3.25]</td>
<td>1.95 [1.67, 2.23]</td>
<td>&lt;.001</td>
<td>.06</td>
</tr>
<tr>
<td>10. Now consider danger to other people in the simulation. Assuming that this simulation was real, rate how much danger you felt that other people in the were in at the point (described above) that you perceived the most danger in this simulation.</td>
<td>6.36 [6.11, 6.60]</td>
<td>5.68 [5.43, 5.93]</td>
<td>6.06 [5.82, 6.29]</td>
<td>&lt;.001</td>
<td>.04</td>
</tr>
<tr>
<td>11. Knowing that this was a use-of-force training session in which officers are likely to make some mistakes, rate how justified you think (you/the officer) were in the level of force that (you/the officer) chose to use in this simulation.</td>
<td>6.94 [6.71, 7.17]</td>
<td>5.83 [5.60, 6.06]</td>
<td>6.83 [6.65, 7.00]</td>
<td>&lt;.001</td>
<td>.19</td>
</tr>
</tbody>
</table>

**Note.** * p < .05, ** p < .01, *** p < .001. Significance tests are based on 2-tailed, one-way ANOVAs with post-hoc independent groups pair-wise t-tests using Holm-Bonferroni correction. The values for p and $\eta^2_p$ in the right-hand columns are for the one-way ANOVA. For questions 8 – 11, responses were rated on a 1 (lowest level) – 7 (highest level) Likert scale.
(1) At what point did (you/the officer) first draw (your/their) gun?

The One-Way ANOVA performed on the mean response to this question in seconds yielded a nonsignificant main effect of group, $F(2, 371) = 0.16$. This is consistent with results from Experiment 1.

(2) At what point did (you/the officer) first shoot (your/their) gun?

The One-Way ANOVA performed on the mean response accuracy to this question yielded a nonsignificant main effect of group, $F(2, 385) = 0.24$. This is consistent with results from Experiment 1.

(3) How many times did (you/the officer) shoot (your/their) gun?

The mean number of shots fired by officers in the simulator was 3.66 [3.30, 4.01]. The One-Way ANOVA performed on the estimated mean number of shots fired yielded a significant main effect of group, $F(2, 411) = 7.69$, $p < .001$, $\eta^2_p = .04$. As can be seen in Table 3, officers who participated online estimated that the number of shots fired was greater than estimates by civilians $t(411) = 2.27$, $p = .047$, $d = .27$, suggesting a cohort effect. The difference between estimates of officers who participated live versus online was not significant, $t(411) = 1.63$, $p = .104$, suggesting no effect of presentation medium. Consistent with the results of Experiment 1, estimates of the number of shots fired were lower for civilians than for both groups of officers.

(4) Which equipment did (you/the officer) pull out first?

The coding of responses to this question was binary, accurate or not accurate. The One-Way ANOVA performed on the mean response accuracy to this question yielded a significant main effect of group, $F(2, 411) = 9.70$, $p = .001$, $\eta^2_p = .05$. As can be seen in Table 3, officers who participated live were most accurate recalling which equipment they had pulled out first. Responses of officers who participated online did not differ from those of civilians $t(411) = 0.90$, $p = .371$, suggesting no effect of cohort. The difference between estimates of officers who participated live versus online was significant, $t(411) = 4.11$, $p < .001$, $d = .50$, suggesting an effect of presentation medium, with officers who participated live being more accurate than those who participated online. As in Experiment 1, reports of which equipment was pulled out first were most accurate for officers who participated live.

(5) What was the total duration of this event?

The total duration of each event was approximately 60 s, with some variability depending on the action taken by the officer at the end of each scenario. As can be seen in Table 3, participants in all three groups over-estimated the duration of the scenarios. The One-Way ANOVA performed on the mean estimates of the duration of the event yielded a significant main effect of group, $F(2, 411) = 12.70$, $p < .001$, $\eta^2_p = .06$. Estimates of officers who participated online were significantly more accurate than those of civilians,
$t(411) = 4.63, p < .001, d = .56$, suggesting an effect of cohort. In addition, estimates of officers who participated online were significantly more accurate than those of officers who participated live, $t(411) = 4.04, p < .001, d = .49$, suggesting an effect of presentation medium. Perhaps the higher level of stress for officers participating live than online led to their over-estimation of the duration of the event.

**State of mind questions**

*Open-ended state of mind questions.* This set of analyses assessed responses to the two open-ended recall questions about the officers’ state of mind when they participated in the training simulator. The same method for coding these two questions was used in Experiments 1 and 2.

(6) **Assuming that this simulation was real, in one sentence specify the point at which you started to feel that (you/the officer) were in danger in this simulation.**

The One-Way ANOVA performed on these data yielded a significant main effect of group, $F(2, 401) = 9.18, p < .001, \eta^2_p = .04$. As can be seen in Table 3, officers who participated online perceived the start of danger significantly sooner than civilians, $t(401) = 4.13, p < .001, d = .50$, suggesting an effect of cohort. However, estimates of the start of danger did not significantly differ between officers who participated online and those who participated live, $t(401) = 1.06, p = .288$, suggesting no effect of presentation medium on this question.

(7) **Assuming that this simulation was real, in one sentence specify the point at which you felt that (you/the officer) were in the most danger in this simulation.**

The One-Way ANOVA performed on these data yielded a nonsignificant difference among groups, $F(2, 400) = 2.76, p = .065$. Interestingly though, although both groups of officers had indicated that they perceived the start of danger to be earlier in the scenarios than did civilians (responses to question 6 above), officers indicated in their responses to question 7 that they perceived peak danger at about the same time in the scenarios as did civilians; officers detected cues for start of danger earlier in the incident than did civilians but detected peak danger at the same point in time.

*Likert scale state of mind questions.* Four Likert scale questions assessed participants’ state of mind at various points throughout each simulation scenario. Again, each question was responded to on a 1 (lowest level) – 7 (highest level) Likert scale. For each question, officers who participated live were responding from their own point of view when they were participating in each incident in the simulator; civilians and the online officers were asked to respond from the point of view of the participating officer in the BWC recordings they were assigned to view.

Responses to the 4 Likert scale questions about the officers’ state of mind were first analyzed with a One-Way Multivariate Analysis of Variance (MANOVA), comparing overall performance across the 3 groups. The MANOVA revealed that across these 4 questions, responses significantly differed among the three groups, $F(8, 818) = 15.60, p < .001; \text{Pillai’s Trace } = .26, \eta^2_p = .13$. 

K. PEZDEK ET AL.
(8) **Rate how much danger you felt (you/the officer) were in at the point that (you/they) felt the most danger in this simulation.**

As presented in Table 3, participants in all three groups responded with a high rate of perceived danger, as would be predicted in real-world use-of-force incidents, especially those involving domestic violence. This suggests that both officers and civilians in this study found the simulations immersive and were highly engaged. The One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 3.07, p = .048, \eta^2_p = .02$. The officers who participated online perceived that the incidents were significantly more dangerous than the civilians, $t(411) = 2.48, p = .041, d = 0.30$, suggesting a significant effect of cohort. Ratings of peak danger did not significantly differ between officers who participated online and those who participated live, $t(411) = 1.24, p = .433$, suggesting no effect of medium.

(9) **Rate how much control of the situation (you/the officer) felt (you/they) had at the point that (you/they) felt the most danger in this simulation.**

As can be seen in Table 3, participants in all three groups responded with a low rate of perceived control, as would be predicted in a real-world use-of-force incident. The One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 12.90, p < .001, \eta^2_p = .06$. Officers who participated online perceived that the officer in the BWC video had significantly less control than did both the civilians, $t(411) = 5.08, p < .001, d = 0.61$, and the officers who participated live, $t(411) = 2.65, p = .017, d = 0.32$, suggesting significant effects of both cohort and medium.

(10) **Now consider danger to other people in the simulation. Assuming that this simulation was real, rate how much danger you felt that other people in the video were in at the point (described above) that you perceived the most danger in this simulation.**

As can be seen in Table 3, participants in all three groups responded with a high rate of perceived danger to other people in the videos, again indicating that the officers and the civilians in this study responded to the simulations as would be predicted in a real-world use-of-force incident. The One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 7.90, p < .001, \eta^2_p = .04$. The only significant pair-wise difference was between the two groups from Experiment 1, officers who participated live and civilians, $t(411) = 3.97, p < .001, d = 0.48$.

(11) **Knowing that this was a use-of-force training session in which officers are likely to make some mistakes, rate how justified you think (you/the officer) were in the level of force that (you/the officer) chose to use in this simulation.**

As can be seen in Table 3, and consistent with the pattern of results in Experiment 1, participants in all three groups responded with a high rate of perceived justification for the level of force used. However, the level of justification reported by both groups of officers revealed a ceiling effect; the officers almost always reported that they were maximally justified (i.e. provided a rating of 7) in the level of force they had used. Nonetheless,
the One-Way ANOVA performed on the mean Likert scale responses to this question yielded a significant main effect of group, $F(2, 411) = 46.96, p < .001$, $\eta^2_p = .19$. Ratings of the level of justification were significantly higher for officers who participated online than for civilians, $t(411) = 7.89, p < .001$, $d = 0.95$, suggesting a significant effect of cohort. However, the difference in ratings of justification between the officers who participated live versus online was not significant, $t(411) = .92, p = .357$, suggesting no effect of medium.

In Experiment 2, of the 11 test items, responses on 7 questions revealed a significant effect of cohort reflecting differences between officers and civilians, responses on 3 questions revealed a significant effect of medium reflecting differences between viewing the incident live versus online, and of these 7 test items, responses on 2 questions revealed a significant effect of both cohort and medium.

**General discussion**

The primary goal of this study was to determine if there are differences between the perspective of an officer who participated in a use-of-force incident and civilians who later only viewed the officer’s BWC footage of the incident, and yes, there were notable differences evident in the results of Experiment 1. Compared to civilians, officers were more accurate remembering the number of shots fired (civilians underestimated the number of shots fired) and which equipment they first pulled out. Officers rated the incidents as more dangerous to themselves and to other people, indicated that they had less control during the incident, and perceived the start of danger earlier in the incident. Perhaps this is why officers indicated that they were more justified in the level of force that had been utilized. Critically, these results suggest that in assessing the actions of law enforcement officers in use-of-force incidents, civilians (including triers of fact) are not likely to be using the same standards used by officers, thus distorting the application of the principal legal ruling in *Graham v. Connor* (1989).

Four mechanisms are considered to account for these differences: (a) degree of schematic processing related to information in the dispatch call, (b) differences in presentation medium between processing a live event versus a BWC recording of it, (c) higher stress at encoding producing more memory impairments and perceptual distortions by officers, and (d) cohort differences between officers and civilians based on expectations and perspective derived from officers’ training and experience. The lack of significant results as a function of the dispatch call variable in Experiment 1 leads us to reject the first proposed mechanism. The other three mechanisms remain candidates.

The role of presentation medium – processing a live event versus a BWC recording of it – was examined in Experiment 2 by comparing responses of officers who participated live versus online. This comparison was significant for only 3 of the 11 questions:

4. *Which equipment did (you/the officer) pull out first?*
5. *What was the total duration of this event?*
9. *Rate how much control of the situation (you/the officer) felt (you/they) had at the point that (you/they) felt the most danger in this simulation.*
Further, the pattern of results on these 3 questions was not consistent, thus failing to support the effect of medium as a mechanism to account for the response differences in Experiment 1 between officers who participated live in the shooting simulation and civilians who participated online and only viewed the BWC recording of an officer’s live participation. Referring to the 3 questions above, live officers were most accurate reporting which equipment they had pulled out first, online officers were most accurate estimating the duration of the event, and live officers perceived more control in the event.

The above comparisons between the pattern of responses for officers who participated live versus online also presents a test of the stress mechanism to account for the results in Experiment 1. Officers who participated live would be predicted to have experienced a higher level of stress than those who participated online. This is consistent with results reported by Hope et al. (2016) in which police officers serving as active witnesses to a stressful use-of-force simulation exhibited higher measures of heart rate and heart rate variability — reflecting higher levels of stress — than officers serving as observer witnesses. Nonetheless, as indicated above, in Experiment 2, for 8 of the 11 questions, there was no significant differences in responses between these two groups of officers, failing to support the stress mechanism to account for the results of Experiment 1.

Among the mechanisms proposed in this study to account for differences between perceptions of a use-of-force incident by officers and civilians, the most compelling is cohort differences between officers and civilians, cohort differences likely to be based on differences in expectations and perspective derived from officers’ training and experience. Although research findings suggest that police officers do not have superior eyewitness memory than civilians (for a review of this research see Pezdek & Reisberg, 2022; Reisberg & Pezdek, in press), based on their training and experience, police officers’ focus of attention is more likely to be guided by their beliefs and expectations (for classic data, see Yarbus, 1967). Although police officers are trained to have a decision-making mental model guiding their response to a critical incident, civilians view BWC footage without this framework. On that basis, it is plausible that officers will provide more useful and more accurate accounts of what they saw at a crime scene — not because they saw more, or had better memory per se — but, instead, simply because they knew where to focus their attention.

It is also worth noting that in both experiments, officers provided very high ratings of their justification in the level of force used. This was consistent for officers who participated live and online. The level of force used by officers, measured in terms of what equipment they drew first, when they first drew their gun, when they first shot their gun, and how many shots were fired, varied across officers. Regardless, officers felt highly justified in the level of force they had chosen to use. If this finding generalizes to officers’ state of mind during real-world use-of-force incidents, it suggests that although an event may be fast moving, complex, and unpredictable, officers consider that the level of force with which they responded reflects their best judgment regarding what the situation required.

One might question whether participants in this study were truly engaged in the simulation and whether their state of mind approximated that in a real-world use-of-force incident. However, analyses of the Likert scale data for both experiments suggest that our participants found the simulations immersive and were highly engaged. Both officers and civilians reported high ratings of perceived danger and low ratings of perceived
control, the precise direction of effects that would be predicted in a real-world use-of-force incident.

There are a few limitations to this study. First, because we only used 2 training scenarios – both use-of-force incidents with officer involved shooting incidents of domestic violence – there is a potential ‘stimulus sampling’ issue. It is not clear that the use of only 2 training scenarios would permit generalization of these results to other types of use-of-force incidents. That said, it is fairly common for researchers to use only two videos to assess cognitive processing of crime videos.

A second limitation is that the civilians in our study who viewed an officer’s BWC video did so only once. This was necessary in our study to make unconfounded comparisons between officers (for whom their live participation was only possible once) and civilians. We recognize, however, that jurors are often permitted to view an officer’s BWC video more than once, although the number of repetitions is typically restricted in the courtroom context. Jurors aside, civilians viewing videos of high-profile crimes have the opportunity to view the videos multiple times. It is possible that with repeated viewing of the video, responses to viewing the incident online may more closely approximate responses of officers experiencing the event live. Future research will be needed to test this hypothesis.

In its ruling in *Graham v. Connor* (1989), the US Supreme Court defined the ‘reasonableness criterion’ used to assess the actions of law enforcement officers in use-of-force incidents, as being ‘judged from the perspective of a reasonable officer on the scene, rather than with the 20/20 vision of hindsight’ (*Graham v. Connor*, 1989). Our results suggest that in assessing the perspective of an officer, it is important to acknowledge that this perspective is likely to differ from that of civilians (including triers of fact) who only view the officer’s BWC recording afterward. And, this difference appears to result from expectations and perspectives derived from officers’ extensive training and experience.

We end with a suggestion for bringing officer and civilian perspectives into better alignment when making judgments about reasonableness, especially in the courtroom; that is, the use of expert witnesses. Here, the expert testimony might be from a police instructor who could educate the jury about relevant legal standards and what a reasonable officer should do based on their training. This suggestion follows from research that has shown that to some extent, people can set aside their own viewpoint and adopt the mindset of someone else, if they are informed about the perspective of the other person (Eyal et al., 2018; Zhou et al., 2017). In fact, similar to the matter at hand, Vardsveen and Wiener (2022) reported that using expert testimony to educate jurors about officers’ training and experience helped mock jurors adopt the perspective of a reasonable officer in assessing a crime scenario and moderated the effect of extralegal factors (including suspect’s race) on legal judgments. The use of expert testimony, combined with jury instructions regarding the ‘reasonable officer standard’ of *Graham v. Connor* (1989), are likely to better align the perspectives of jurors with those of police officers in judging whether the force used by an officer was justified. However, we do not want to underestimate the challenge of this task. Given that in our study officers’ perspectives had been altered by an average of more than 15 years of training and experience, it is unlikely that civilians can adopt the perspective of ‘a reasonable officer on the scene’ without quite a significant intervention.
Notes

1. Shane et al. (2017) reported that fatal police shootings in the U.S. in 2015–2016, although always unfortunate, are relatively uncommon, averaging about 0.336 per 100,000 citizens and are relatively stable between these two years.

2. For the purposes of this study, we consider civilians to be anyone except the participating officers; this includes members of the public, jurors, legal professionals.

3. For a review of the psychological research on the use of BWCs, see Pezdek (2022).

4. While the Federal ‘reasonableness standard’ established by Graham is the minimum protection guaranteed, states may impose a higher standard.

5. Marr et al. (2021) suggested that the effects of stress may differ between tasks typical of fundamental memory studies and those typical of eyewitness memory studies. However, this view seems at odds with the empirical evidence.

6. There are separate issues that concern the credibility of video evidence in general and also whether triers of fact are likely to rely more on an officer’s account or civilians’ interpretation of video evidence. These issues are different than those considered in this study but have been discussed elsewhere. See for example Jones et al. (2017) and Pezdek and Reisberg (2022).

7. The COVID-19 pandemic necessitated that the civilian sample participate on MTurk rather than in-person.

8. This study was preregistered on OSF as a 2-groups design, referring only to the two groups of civilians. Comparisons of the civilian participants with police officers were then included as exploratory analyses. This is because OSF has strict guidelines for the use of previously collected data. Therefore, although conducting an analysis that compared the previously collected officer data with the new civilian data was the most conceptually appropriate approach, this needed to be pre-registered as an exploratory analysis by OSF standards.


10. See: https://osf.io/dk253/?view_only=7c4c751bff86487aa80e3050a36be747

11. In the study by Pezdek et al. (2022), the officers responded to the same questions later at two points in time. Importantly, in that study, the officers had viewed their own BWC video prior to answering the subsequent questions for one scenario but not for the other scenario. In the current study, none of the officers had viewed their BWC video prior to answering the Time 1 test questions.

12. Based on a report from the FBI’s Law Enforcement Officers Killed and Assaulted database, 8.5% of all officers killed on duty between 2011 and 2020 were responding to domestic disturbance or domestic violence calls. See: https://www.fbi.gov/services/cjis/ucr/leoka

13. In this analysis, and all analyses in this study, across test questions different numbers of participants were excluded because of missing data, primarily related to some ambiguous responses that could not be coded.

14. For all statistical results see: https://osf.io/dk253/?view_only=7c4c751bff86487aa80e3050a36be747

15. Throughout, 95% confidence intervals are provided in brackets.

16. This includes question 8 for which responses of officers significantly differed from the civilians without the dispatch call but not for civilians with the dispatch call.


18. All analyses involving police officer samples (in-person or online officers) were, strictly speaking, exploratory analyses for the reasons discussed in Footnote 6.

19. As a reminder, in this analysis, and all analyses in this study, across test questions different numbers of participants were excluded because of missing data, primarily related some ambiguous responses that could not be coded.

20. Pair-wise statistical comparisons between officers who participated live and civilians with dispatch call will not be reported here because this contrast was reported for these same conditions in Experiment 1.

21. For all statistical results see: https://osf.io/dk253/?view_only=7c4c751bff86487aa80e3050a36be747
22. 95% confidence intervals are provided in brackets.
23. We did not, however, assess physiological levels of stress experienced by officers in this study.

Author contributions
KP developed the study concept and obtained funding for this research. KP and TS drafted the research design. TS and JB collected the data from the live police officers. TS built the Qualtrics survey and collected the MTurk data from civilians and online officers. All authors participated in analyzing and interpreting the data. KP drafted the manuscript, and TS participated in the revision process. All authors read and approved the final manuscript.

Author note
This research was funded by a grant to Kathy Pezdek from the National Science Foundation, Program in Law and Social Sciences. We have complied with APA ethical standards in the treatment of subjects in this study, and these data have not previously been disseminated or published. The authors report there are no competing interests to declare.

This experiment was formally preregistered on Open Science Framework and can be accessed at https://osf.io/dk253/?view_only=7c4c751bff86487aa80e3050a36be747. All anonymized data are publicly available on OSF.

We are grateful to Erik Fritsvold, University of San Diego, Michael Scott, Arizona State University, and Seth Stoughton, University of South Carolina Law School, for helping us recruit the officers who participated in Experiment 2.

Disclosure statement
No potential conflict of interest was reported by the author(s).

ORCID
Kathy Pezdek http://orcid.org/0000-0001-8602-5142
Tyler Shapland http://orcid.org/0000-0001-6581-3879
Jessica Barragan http://orcid.org/0000-0003-2331-649X

References


