

## RESEARCH ARTICLE

# Body-Worn Camera Versus Officer Reports: How Inconsistencies Influence Public Perceptions of Trust, Accuracy, and Punishment

Kris-Ann S. Anderson<sup>1,2</sup>  | Kristyn Jones<sup>1,2</sup> | Deryn Strange<sup>1,2</sup>

<sup>1</sup>Department of Psychology, Graduate Center, City University of New York, New York, New York, USA | <sup>2</sup>Department of Psychology, John Jay College of Criminal Justice, City University of New York, New York, New York, USA

**Correspondence:** Kris-Ann S. Anderson ([kranderson@jjay.cuny.edu](mailto:kranderson@jjay.cuny.edu))

**Received:** 29 April 2025 | **Revised:** 19 August 2025 | **Accepted:** 28 August 2025

**Funding:** The authors received no specific funding for this work.

**Keywords:** body-worn cameras | credibility | evidence consistency | police reports | use of force

## ABSTRACT

Some officers advocate for watching their body-worn camera (BWC) footage before writing incident reports, fearing inconsistencies may damage their credibility. Across two experiments ( $N = 1580$ ), we examined how inconsistencies influence public perceptions of trust, accuracy, and punishment decisions for an officer involved use-of-force incident. In Experiment 1, participants read a brief description and the officer's report but did not see BWC footage; in Experiment 2, participants read the officer's report and watched ambiguous BWC footage. We manipulated whether the report was described as consistent or inconsistent with the footage and whether the officer wrote the report before or after watching the footage. Overall, inconsistency reduced trust and accuracy ratings and, in Experiment 1, increased charging decisions and guilt judgments. Additionally, participants preferred officers watch BWC footage after writing their reports. These results suggest a need for public education about interpreting inconsistencies and that officers have reason to fear inconsistencies.

In 2017, Miguel Richards was fatally shot by police officers during a wellness check. The officers reported that Miguel had a knife and a firearm, justifying their actions (Lopez 2023). At the time, New York City had mandated that all officers wear body-worn cameras (BWCs). However, when the department released the BWC footage, the recording was ambiguous. In the footage, Miguel is standing in the corner of his bedroom, but it is unclear whether he is holding a firearm, a knife, both, or nothing at all. This ambiguity invited the public to question the officers' accounts. One person stated, "...he was not threatening them... There's no reason to use lethal force" (Lowenkron 2021 paras. 8 and 23). But Miguel's case is far from an unusual situation; shaky, low-quality video and frequent obstructions mean BWC footage is often ambiguous and that footage is often released. Thus, we wondered, how do apparent inconsistencies between what an officer said happened and what the footage shows

influence public perceptions? Would people be less willing to trust the officers' accounts?

## 1 | Police BWCs and Policies

BWCs experienced a widespread push in 2014 after police-involved shootings sparked national protests (Benner 2019; Hermann and Weiner 2014). Over the past decade, federal financial support for BWCs has increased by over \$83 million (White House 2024). The most recent report estimated that 60% of local law enforcement had implemented BWCs in 2016 (Hyland 2018); this number has likely increased as federal policies have called for broader implementation (House 2022; Department of Homeland Security 2023). The overarching support comes from the belief that BWCs provide an objective record of police interactions and

should improve trust between officers and community members by promoting transparency and accountability (Chapman 2019; Office of Justice Programs 2016; Hutto and Green 2016). But have BWCs lived up to the hype?

Certainly, BWC footage has been at the center of some high-profile cases. It has both justified (Cullen 2018; Yang 2019) and vilified officers' actions (Dimmick 2022). However, current evidence shows no consistent improvement in police behavior for use-of-force incidents, pedestrian stops, or arrests (Lum et al. 2020). What about the promise of transparency? Here, the evidence is more mixed. On the one hand, the public supports the release of BWC footage (Todak et al. 2021). But on the other hand, both recent media reports and public surveys reveal that, for many, the desired transparency has not been realized (Lopez 2023; Tavernise 2024; Farooq 2023; Umansky 2023; Wright and Headley 2021).

We suspect the discontent stems from the policies governing the use of BWCs. For example, many departments allow cameras to be turned on and off by the officer, which can result in critical moments not being recorded (Mustian 2018). Additionally, there are concerns about the public's access to BWC footage. When footage does exist, it is not always released (Farooq 2023) or is not released in a timely manner (Lopez 2023; Umansky 2023). These practices appear to hamper the public's perception of transparency, especially given that a survey showed 59% support immediately releasing the footage (Todak et al. 2021). But, perhaps the greatest challenge to the perception of BWCs is that police are typically able to review their footage before writing a report.

Termed "pre-report review," researchers, legal scholars, and civil liberty groups are united in their objections to this policy. They are concerned that pre-report review undermines the independence of the evidence and may alter officers' descriptions of the scene, the sequence of their actions, and ultimately, their memories of the incident (Jones and Strange 2022; Pezdek 2015, 2022). Indeed, one study revealed that when people review BWC footage, they change their reports to only include the details shown on camera (Adams et al. 2020), resulting in reports that summarize the footage rather than offer a unique source of information (Adams et al. 2020; Grady et al. 2016; Vredeveltdt et al. 2021). Do these perfectly matched reports make people suspicious of a cover-up and, thus, question an officer's credibility?

Police do not think so. Instead, police departments say pre-report review is necessary *precisely* because it produces statements consistent with BWC footage. From the police department's perspective, the consistency between reports and BWC footage will promote trust and ensure accurate recollections without relying on the frailties of human memory. Put differently, departments understand that there could (perhaps even should) be inconsistencies between BWC footage and police reports due to how people encode and remember events. But some fear the public will not tolerate the idea that police could simply misremember something, thus undermining their credibility (Miller et al. 2014).

The departments' concern is understandable; after all, officers have used suspect statement consistency as a measure

of credibility and truthfulness for decades (e.g., Akehurst et al. 1996; Deeb et al. 2018; Fisher et al. 2009). They are not the only ones. Lawyers, officers, and mock jurors use consistency to measure witness credibility (Brewer et al. 1999). Moreover, jurors believe inconsistent statements are less accurate and reliable than consistent ones, regardless of the witness' age (Brewer et al. 1999; Berman et al. 1995; Granhag and Strömwall 2000). In fact, jurors are less likely to convict defendants when eyewitness statements are inconsistent (Palmer et al. 2016). These findings highlight how consistency across statements is commonly interpreted as a measure of credibility and truthfulness, whereas inconsistencies lead to perceptions of dishonesty.

The previous studies, however, have examined the consistency of a single individual over the course of their statements and testimony. What we do not know is how people perceive inconsistencies between different types of evidence, such as an officer's report and ambiguous BWC footage—especially in use-of-force cases. Had there been a criminal trial in the Miguel Richards' case, jurors would have been instructed to apply the *Graham v. Connor* (1989) standard<sup>1</sup> to determine if the officers' use of force was "objectively reasonable". That is, jurors would have been instructed that the officer's actions should be judged from the perspective of a reasonable officer on the scene while considering the facts, without the benefit of hindsight. But no jurors considered that standard in Miguel Richards' case; instead, a judge in a civil lawsuit determined the officers' actions were justified (*Estate of Miguel Antonio Richards v. City of New York* 2024). Yet, in the court of public opinion, the inconsistencies between what the officers said happened and what the footage showed had already sparked public debates (e.g., Lopez 2023; Lowenkron 2021).

Because most use of force cases do not proceed to criminal trial (Stinson and Wentzlof 2019), we did not call for participants to apply the *Graham* standard. We, instead, focus on the court of public opinion, where public perceptions of officers can influence community relations and drive changes in policy. Specifically, we ask when BWC footage does not confirm or deny what the officer said happened, how do people interpret discrepancies? Does the ambiguous footage influence public perceptions of officer trust, or do people give the officer the benefit of the doubt? Such questions are critical to current BWC policy discussions.

## 2 | Overview of Studies

We conducted two experiments to determine people's perceptions of an officer's use of force against a civilian. We manipulated whether the report was described as consistent or inconsistent and whether the officer wrote the report before or after watching the BWC footage. We had several questions to answer: (1) Are people sensitive to the potential consequences of pre-report review? (2) How does pre-report review affect perceptions of the officer? (3) What happens when people are told the officer's report is inconsistent with BWC footage, but don't see those inconsistencies? (Experiment 1) (4) What happens when people see ambiguous BWC footage and are told the footage is inconsistent with the officer's report? (Experiment 2).

We predicted that participants would perceive the officer's report as less accurate and trustworthy when the officer engaged in pre-report review and when the report was described as inconsistent with the footage. We also expected participants to rate BWC footage as more accurate and trustworthy than the officer's report, with the preference becoming more pronounced when they are told about an inconsistency. Additionally, we predicted that officer punishment decisions would be influenced by the report being described as consistent or inconsistent. Finally, we expected that stronger identifications with the police would be associated with higher trust and accuracy ratings of the police report, regardless of consistency.

### 3 | Experiment 1

#### 3.1 | Design

We conducted a 2 (BWC Review: before, after)  $\times$  2 (Consistency: consistent, inconsistent)  $\times$  2 (Evidence Type: report, video) mixed design with BWC Review and Consistency as between-subjects factors and Evidence Type as a within-subjects factor. We conducted an a priori power analysis to determine the sample size needed for a medium-sized interaction ( $f = 0.247$ ) for the two between-subjects factors: 2 (BWC Review: before, after)  $\times$  2 (Consistency: consistent, inconsistent). The results showed that 1000 participants were needed to provide 89% power at  $\alpha = 0.05$ . We did not power the study for the within-subjects factor as mixed designs generally have greater statistical power and require fewer participants (Brysbaert 2019; Lakens 2022).

#### 3.2 | Participants

We recruited 1756 United States (U.S.) participants through Amazon Mechanical Turk (MTurk; Litman et al. 2017), adding additional safeguards to improve data quality.<sup>2</sup> The average completion time was 19 min. We excluded 384 participants who failed at least one of the two manipulation checks for a final dataset of 1372 (57.9% male, 48.7% female, 0.4% non-binary) with a mean age of 40.60 (SD = 13.19). Overall, participants were more highly educated than the average U.S. population, with 57% earning at least a bachelor's degree or higher, compared to the 36% nationally in 2022 (McElrath 2025; see Table 1 for demographics).

#### 3.3 | Materials

##### 3.3.1 | Report

We created a 316-word mock officer's report that described a non-fatal use of force incident between Officer O'Brien and a civilian. In the report, Officer O'Brien describes being dispatched to a location and attempting to engage with a young man, Mr. Perkins. However, when Perkins did not comply with the officer's commands, the officer drew his weapon. Officer O'Brien reports he believed the civilian had a knife and fired, shooting Perkins in the leg. The report ends by explaining the civilian received medical attention.

**TABLE 1** | Demographic characteristics of participants in Experiments 1 and 2.

	Experiment 1	Experiment 2
	%	%
Race/Ethnicity		
White	75.29	69.71
Black or African American	11.37	14.42
Hispanic or Latino	5.54	7.21
Asian	5.69	4.33
Hawaiian or Pacific Islander	0.29	0
Other	1.82	4.33
Education		
Bachelor's	41.69	35.10
High school diploma	29.15	36.06
Master's or Ph.D.	15.73	14.9
Associate's	12.68	13.46
Did not finish high school	0.73	0.48

Note: Experiment 1:  $N = 1372$  and Experiment 2:  $N = 208$ .

#### 3.4 | Measures

Participants responded to the measures using a 1 (*not at all*) to 7 (*completely*) Likert scale unless otherwise indicated. Due to the similarity of some questions, we highlighted the key differences using bold text in the survey. We italicize those same sections below.

##### 3.4.1 | Trust in Report and Footage

There were two questions related to trust: "How much do you trust that the *officer's statement described what actually happened?*" and "How much do you trust that the *body camera footage showed what actually happened?*"

##### 3.4.2 | Accuracy of Report and Footage

Participants answered two questions to determine their beliefs that both evidence types matched the encounter: "To what extent does the *body camera footage match the actual encounter?*" and "To what extent does the *officer's statement match the actual encounter?*"

Another two questions determined people's beliefs that the evidence types matched what the officer saw happen: "To what extent does the *officer's statement match what the officer saw happen?*" and "To what extent does the *body camera footage match what the officer saw happen?*"<sup>3</sup>

### 3.4.3 | Punishment Decisions

Participants responded to three punishment questions. They rated the extent to which “The officer should be charged and tried in court for his actions in the encounter?” on a 1 (*strongly disagree*) to 7 (*strongly agree*) Likert Scale. Next, they answered, “If the officer is put on trial for his actions in the encounter, should he be found guilty?” and “When making punishment decisions, which piece of evidence should jurors take into consideration the most?” with “what the officer said happened” and “what the body camera footage showed” as response options.

### 3.4.4 | Pre-Report Review

Participants rated their views of pre-report policies to determine public perceptions of this police policy on a 7-point Likert Scale ranging from 1 (not *reasonable at all*), 4 (somewhat *reasonable*) to 7 (*completely reasonable*). The question read, “How reasonable was it that the officer was [wasn’t] able to review his body camera footage before writing his statement?” based on their experimental condition.

### 3.4.5 | Identification With Police Scale (IPS)

We used the shortened 7-item version described in Granot et al. (2014;  $\alpha = 0.87$ ) to measure the extent to which people share similar values with the police on a 7-point Likert scale (1 = *strong disagreement*, 4 = *neither disagree nor agree*, 7 = *strong agreement*; Tyler and Fagan 2008). One example question reads, “If you talked to most police officers, you think you would find that they have similar views to your own on many issues.” Higher scores indicate greater identification with the police.

### 3.4.6 | Officer’s Experience

Participants rated the extent to which they believed the officer was being truthful: “The officer said that the civilian looked like he was going to lunge at him. How much do you believe that this is an accurate representation of what the officer experienced?” We also directly asked participants, “To what extent do you think the officer is telling the truth about what he saw happen in the encounter?”<sup>4</sup> Additionally, those in the inconsistent condition ranked the most likely reasons the officer said there was a knife, but none was visible in the video. The options included (1) “The officer thought he saw a knife,” (2) “The officer is lying about what he saw,” (3) “The officer incorrectly remembered seeing a knife,” and (4) “The officer was correct about seeing a knife but the knife wasn’t visible in the video (e.g., because of poor lighting, poor quality of the video, or the angle of the recording, etc.).”

### 3.5 | Manipulation Check

We included two manipulation check questions: “Did the officer write his statement first or watch the body camera video first?” with three options: (1) he wrote his statement first, (2) he wrote his statement after, (3) not sure. The second question read, “The

officer said there was a knife. Was a knife visible in the video?” with answer choices of yes, no, and not sure.

### 3.6 | Procedure

We told participants they would learn about an interaction between Officer O’Brien and a civilian named Ryan Perkins before being asked about the encounter and the credibility of the officer’s statement. Next, participants were randomly assigned to one of four conditions: (1) BWC review before report and consistent ( $n = 339$ ), (2) BWC review before report and inconsistent ( $n = 330$ ), (3) BWC review after report and consistent ( $n = 368$ ), and (4) BWC review after report and inconsistent ( $n = 335$ ). Those in the BWC Review *Before Report* conditions read:

...Per the department procedure, officers who have engaged in any use-of-force are required to provide an account of the encounter after they watch their body camera footage. On the following pages, you will read the officer’s statement that he provided the night of the encounter after he watched his body camera footage.

Those in the BWC Review *After Report* condition read the same description except the underlined word stated “before.” Following the Review manipulation, participants read the mock police report.

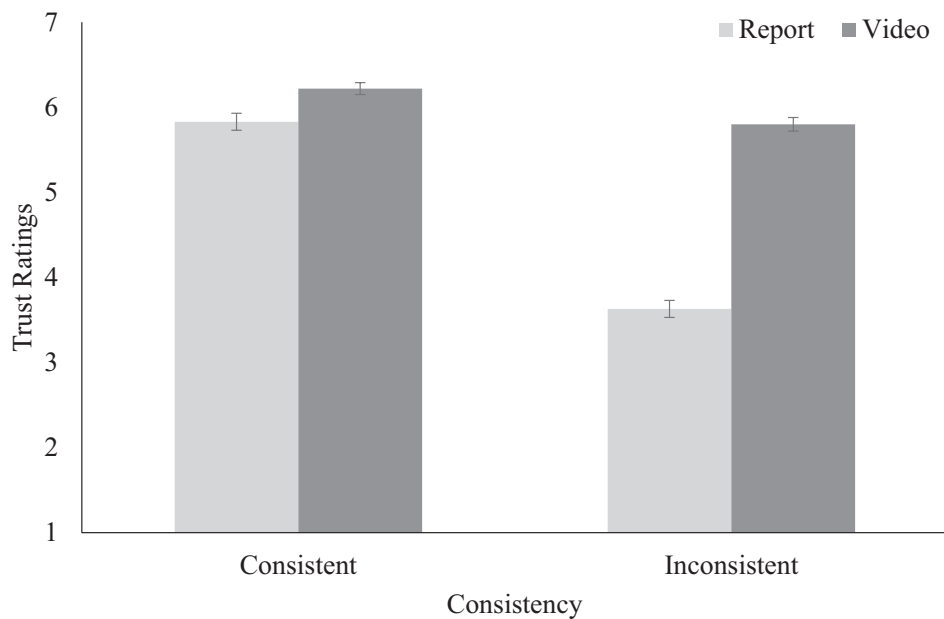
After the report, we repeated the BWC Review manipulation and added the consistency manipulation. Participants in the *Consistent* condition read, “...In their review of the incident, the department found that the officer’s statement was consistent with the body camera footage: The officer’s statement said that he saw a knife and a knife was visible in the video.” Participants in the *Inconsistent* condition read: “...In their review of the incident, the department found that the officer’s statement was inconsistent with the body camera footage: The officer’s statement said that he saw a knife, but no knife was visible in the video.” Finally, participants responded to the dependent measures, manipulation checks, IPS, and demographic questions.

## 4 | Experiment 1 Results

First, we assessed participants’ views on pre-report review policies. An independent samples *t*-test showed participants found it more reasonable for officers to review BWC footage *after* ( $M = 6.08$ ) writing their reports than *before* ( $M = 4.22$ ,  $t(872) = -16.76$ ,  $p < 0.001$ ).<sup>5</sup>

Next, we examined people’s perceptions of the officer’s report and the BWC footage. We conducted 2 (BWC review: before, after)  $\times$  2 (Consistency: consistent, inconsistent)  $\times$  2 (Evidence Type: report, video) mixed Analyses of Variance (ANOVAs). Evidence type was created by using participant responses to the report and body camera footage accuracy and trust questions. For brevity, however, we focus on the 2 (Consistency: consistent, inconsistent)  $\times$  2 (Evidence Type: report, video) results as the Review manipulation did not influence trust, accuracy, or





**FIGURE 1** | Experiment 1: Average perceived trustworthiness of the report and video based on consistency. Error bars represent 95% confidence intervals of cell means. The trust ratings were created as a within-subjects variable from the questions: “How much do you trust that *the officer’s statement* described what actually happened?” and “How much do you trust that *the body camera footage* showed what actually happened?”

punishment ratings. The full  $2 \times 2 \times 2$  ANOVA tables can be found in Tables S1–S3.

#### 4.1 | Trust in Report and Footage

First, did evidence type and consistency influence participant trust ratings? Yes. As shown in Figure 1, the consistency and evidence type interaction was significant such that trust ratings for the video and report depended on consistency,  $F(1, 1368) = 466.48$ ,  $p < 0.001$ . More specifically, participants trusted the BWC footage over the report in both conditions; however, there was a greater difference between the ratings when we told participants that the evidence was inconsistent ( $M_{\text{diff}} = 2.17$ ,  $d = 1.12$ ) versus consistent ( $M_{\text{diff}} = 0.39$ ,  $d = 0.39$ ). Additionally, participants reported overall higher trust ratings in the consistent ( $M = 6.02$ ) versus inconsistent groups ( $M = 4.71$ ). As predicted, participants also indicated the video was more trustworthy compared to the report ( $M_{\text{diff}} = 1.28$ ). In other words, there were significant main effects of consistency ( $F(1, 1368) = 694.31$ ,  $p < 0.001$ ,  $d = 0.92$ ) and evidence type ( $F(1, 1368) = 957.25$ ,  $p < 0.001$ ,  $d = 0.71$ ). All other  $p$  values were not significant ( $p > 0.31$ ; see Table S1).

#### 4.2 | Accuracy of Report and Footage

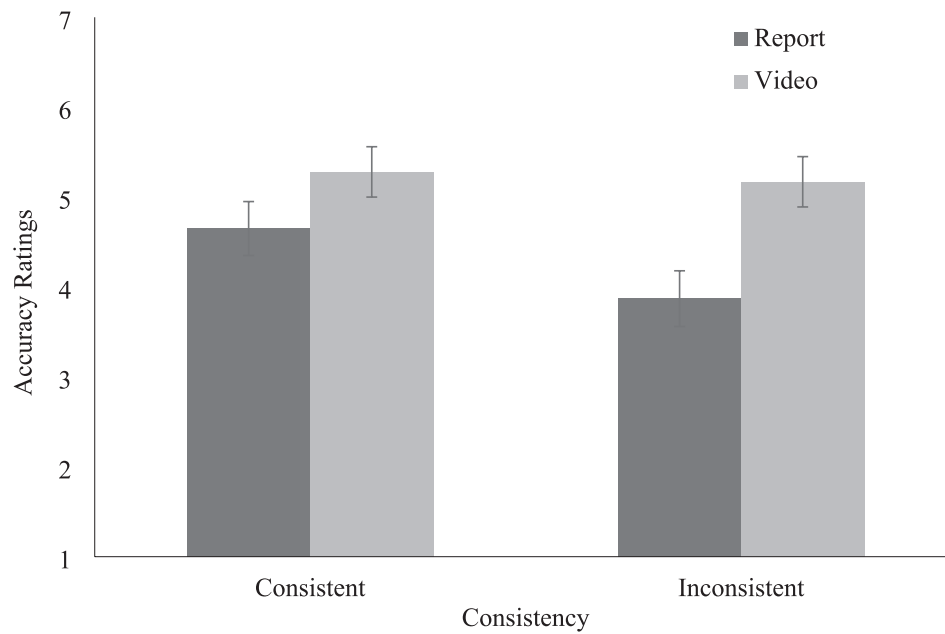
We then examined whether our manipulation impacted participants’ belief that the BWC footage and report represented what happened in the encounter. Here too, the evidence type and consistency interaction was significant,  $F(1, 1368) = 179.20$ ,  $p < 0.001$  (see Figure 2). In particular, participants thought the video was more accurate than the report, but the difference between the ratings was greater for those in the inconsistent ( $M_{\text{diff}} = 1.07$ ,  $d = 0.55$ ) condition, whereas there were no differences for those in the consistent condition ( $M_{\text{diff}} = 0.01$ ,  $d = 0.02$ ).

Additionally, we saw a main effect of consistency where participants in the consistent group provided overall higher accuracy ratings than in the inconsistent group ( $M_{\text{diff}} = 2.02$ ,  $F(1, 1368) = 1134.40$ ,  $p < 0.001$ ,  $d = 1.47$ ). Participants also provided overall higher ratings for the BWC footage compared to the officer’s report ( $M_{\text{diff}} = 0.54$ , reflecting a main effect of evidence type,  $F(1, 1368) = 189.54$ ,  $p < 0.001$ ,  $d = 0.34$ ). All other  $p$  values were not significant ( $p > 0.19$ ; see Table S2).

#### 4.3 | Punishment Decisions and Officer Experience Ratings

We next examined whether participant perceptions influenced judgments about officer punishment. We conducted regressions using BWC Review, Consistency, IPS, Review  $\times$  Consistency, and Consistency  $\times$  IPS as predictors. Overall, the multiple regression model was significant,  $F(5, 1366) = 66.53$ ,  $p < 0.001$  (See Table 2). Here, the effect of inconsistent evidence on officer charging decisions depended on how strongly people identified with police ( $SE = 0.08$ ,  $p < 0.001$ ). Specifically, when the evidence was inconsistent, participants with higher IPS scores were less likely to support charging decisions. Additionally, higher IPS scores were associated with reduced charging decisions overall ( $SE = 0.06$ ,  $p < 0.001$ ). However, the strongest predictor was evidence consistency, with those in the inconsistent group being more likely to suggest that the officer be charged ( $SE = 0.40$ ,  $p < 0.001$ ).

Overall, 68% of participants thought the officer should be found not guilty if he was put on trial for his actions. Using the same predictors, we conducted a logistic regression. Here, the effect of inconsistency on guilt decisions depended on IPS (See Table 3). Specifically, those in the inconsistent condition with higher IPS scores had 50% lower odds of finding the officer guilty. The full model also revealed that when the report and footage were inconsistent, the odds of a guilt judgment were 102 times greater



**FIGURE 2** | Experiment 1: Average beliefs about report and video accuracy based on consistency. Error bars represent 95% confidence intervals of cell means. The accuracy ratings were created as a within-subjects variable from the questions: “To what extent does *the body camera footage* match the actual encounter?” and “To what extent does *the officer’s statement* match the actual encounter?”

**TABLE 2** | Regression results for participants’ support of charging the officer.

Variable	<i>b</i>	<i>b</i> 95% CI	sr <sup>2</sup>	sr <sup>2</sup> 95% CI	Fit
(Intercept)	3.77**	[3.20, 4.33]			
Review (ref. = before)	−0.16	[−0.42, 0.09]	0.00	[−0.00, 0.00]	
Consistency (ref. = consistency)	2.59**	[1.79, 3.38]	0.02	[0.01, 0.04]	
Identification with police	−0.21**	[−0.32, −0.10]	0.01	[−0.00, 0.02]	
Review*consistency	0.21	[−0.16, 0.58]	0.00	[−0.00, 0.00]	
Consistency*Identification with police	−0.27**	[−0.42, −0.12]	0.01	[−0.00, 0.01]	
					$R^2 = 0.196^{**}$
					95% CI [0.16, 0.23]

Note: A significant *b*-weight indicates the semi-partial correlation is also significant. *b* represents unstandardized regression weights. sr<sup>2</sup> represents the semi-partial correlation squared.

Abbreviations: CI = confidence interval; ref. = reference.

\*Indicates  $p < 0.05$ .

\*\*Indicates  $p < 0.01$ .

than when the evidence was consistent. However, the wide confidence interval margins suggested the model may have inflated estimates due to separation of responses. Indeed, when we reviewed guilt judgments by consistency, we found 18.25% of participants in the consistent condition advocated for guilt judgments compared to 46.47% in the inconsistent condition. We then ran a simplified model, removing the interaction terms to assess whether the main effects remained stable. For the simple model, the odds of a guilt judgment were 3.86 times greater in the inconsistent versus consistent conditions ( $p < 0.001$ ). In other words, the participants in the inconsistent condition had a 79.43% probability of finding the officer guilty. And as with the full model, the higher the IPS score, the lower the probability (39.6%) of a guilt judgment, OR = 0.66, 95% CI [0.59, 0.73],  $p < 0.001$ .

Next, we examined participant rankings of the most likely reasons the officer said there was a knife, but none was visible in the video. The majority of the participants ranked “The officer thought he saw a knife” as the first (51.98%) or second (32.28%) most likely reason for the inconsistency. The third most commonly ranked reason was “The officer incorrectly remembered seeing a knife” at 45.5%, and 48.41% ranked “The officer is lying about what he saw.” as the least likely reason for the inconsistency. A full breakdown of the rankings is available in Table S5.

We conducted exploratory ordinal logistic regressions to further explore participant rankings of “The officer thought he saw a knife” and “The officer is lying about what he saw” explanations. IPS significantly predicted the explanation rankings. Specifically, participants who more strongly identified with

**TABLE 3** | Logistic regression results for participants' officer guilt decisions.

Predictor	$\beta$	SE ( $\beta$ )	$z$	$p$	OR	95% CI
Full model <sup>a</sup>						
Intercept	-1.21**	0.42	-2.86	0.004	0.30	[0.13, 0.67]
Review (ref. = before)	-0.15	0.19	-0.79	0.429	0.86	[0.58, 1.26]
Consistency (ref. = consistency)	4.63**	0.59	7.79	<0.001	102.06	[32.50, 333.56]
Identification with police	-0.04	0.08	-0.52	0.602	0.96	[0.82, 1.13]
Review*consistency	0.13	0.26	0.50	0.618	1.13	[0.69, 1.89]
Consistency*Identification with police	-0.70**	0.12	-6.04	<0.001	0.50	[0.40, 0.62]
Simple model <sup>b</sup>						
Intercept	0.53*	0.27	1.96	0.050	1.71	[1.00, 2.92]
Review (ref. = before)	-0.04	0.12	-0.32	0.75	0.96	[0.75, 1.23]
Consistency (ref. = consistency)	1.35**	0.13	10.55	<0.001	3.86	[3.01, 4.97]
Identification with police	-0.42**	0.05	-7.95	<0.001	0.66	[0.59, 0.73]

Abbreviations: CI = confidence interval; OR = odds ratio; ref. = reference; SE = standard error.

<sup>a</sup>This model includes the interaction terms.

<sup>b</sup>This model does not include the interaction terms.

\*Indicates  $p < 0.05$ .

\*\*Indicates  $p < 0.01$ .

police had greater odds of assigning the “thought he saw a knife” explanation as more plausible,  $\beta = -0.29$ ,  $SE = 0.06$ ,  $OR = 0.75$ , 95% CI [0.67, 0.84],  $p < 0.001$ , and greater odds of assigning the “officer is lying” explanation as less plausible ( $\beta = 0.61$ ,  $SE = 0.06$ ,  $OR = 1.85$ , 95% CI [1.64, 2.09],  $p < 0.001$ ).

Finally, we analyzed responses to the question, “When making punishment decisions, which piece of evidence should jurors take into consideration the most?” The majority of participants (86.59%) indicated that jurors should consider the BWC footage when making punishment decisions. Moreover, a logistic regression with IPS as a predictor revealed that for every one-unit increase in IPS scores, the odds of prioritizing BWC footage decreased by 44% ( $\beta = -0.58$ ,  $SE = 0.08$ ,  $OR = 0.56$ ,  $p < 0.001$ ). Put simply, the more strongly someone identifies with police, the more likely they are to say jurors should consider the officer's report over the BWC footage.

#### 4.4 | Experiment 1 Summary

Overall, we found that people consistently rated the BWC footage as more accurate and trustworthy than the officer's report. This effect was magnified when people were told that the two pieces of evidence were inconsistent. Although participants did not see the BWC video themselves, they expressed lower trust in the officer's report when told the report conflicted with the video. These effects extended to punishment decisions. Those in the inconsistent condition were more likely to support charging the officer and delivering a guilt judgment.

However, the effects were not uniform for all participants. People's identification with police (IPS) scores influenced how participants interpreted the evidence. Here, those who more

strongly identified with police were less likely to find inconsistencies as problematic and less likely to support punishment than those who had lower IPS scores. Additionally, those with higher IPS scores were more likely to attribute inconsistencies to a mistake than to deliberate deception.

We also found that, in line with the concerns of legal scholars and researchers, participants found it less reasonable that the officer reviewed the BWC footage before writing their report. However, these views did not significantly influence people's perceptions of trust, accuracy, and punishment decisions. What, then, accounts for this lack of an effect for pre-report review?

One possible explanation is that people may not be educated about the potential consequences of pre-report review and thus do not take issue with these practices. Another is that our vignette said that pre-report review was a department procedure. Thus, knowing the officer saw the BWC footage may not have altered the participants' views because they could have inferred that the officer was just doing their job. Perhaps participants may have inferred intentional deception if the vignette had mentioned that the officer had chosen to watch the video before writing his report. Although this change may have produced a difference in effects, the problem with this scenario is that it is less ecologically valid; pre-report review of footage is typically a department policy instead of an individual officer's choice (Upturn 2017).

Lastly, the review effect could be smaller than we initially hypothesized ( $f = 0.25$ ), and so the study may have lacked the power to detect the proposed interaction. To assess this possibility, we conducted a post hoc sensitivity analysis using G\*Power 3.1.9.6 (Faul et al. 2007), which indicated that our sample of

1372 had 89% power to detect an effect size of  $f=0.09$  at  $\alpha=0.05$ . Therefore, if the effect of BWC review is smaller than  $f=0.09$ , this experiment would not have been able to detect it.

Overall, this experiment simulates a scenario where people do not have access to or do not see the footage themselves, but instead rely on secondhand information that they read in a news article. But what happens when people see the footage for themselves? We explored this question in Experiment 2.

## 5 | Experiment 2

In many real-world cases, BWC footage isn't outright inconsistent; instead, it is ambiguous. Indeed, the camera's physical placement, officers' movement, and other factors can obstruct details and leave the footage open to interpretation. For instance, the Miguel Richards videos do not clearly show whether Richards was holding a weapon (New York City Police Department 2017). In our pre-registered Experiment 2 (<https://osf.io/bygx2>), we sought to replicate and extend Experiment 1. Here, we explore a scenario where participants view ambiguous footage that does not support or contradict the officer's report. Would the suggestion of inconsistencies still influence people's perceptions of the evidence and the officer?

## 6 | Method

We replicated all measures. Our only procedural change was that participants watched a BWC footage clip before reading the officer's report. We also used real-world BWC footage and the corresponding officer's report instead of a mock scenario to increase ecological validity.

### 6.1 | BWC Footage and Report

We used the 600-word incident report obtained from the Athens-Clarke County Police Department (Jones, Crozier, and Strange 2017). The report was modified to state the officer noticed the civilian had a knife in his pocket—mirroring the inconsistency from Experiment 1.

The corresponding 6-min BWC footage, which included audio, shows an interaction between an officer and an intoxicated male civilian. When the civilian does not comply with the officer's demands, the officer hits the civilian with his baton. The video ends with the civilian on the ground while the officer shouts commands for the civilian to put his hand behind his back. Importantly, no knife was visible in the video, creating ambiguity. That is, the video did not support or contradict the officer's claim, leaving room to interpret the officer's actions. Notably, in pilot testing, participants ( $n=36$ ) rated the officer's actions as neither clearly justified nor unjustified ( $M=3.6$  on a 1–7 scale where 4 = *ambiguous*), further suggesting the video was open to multiple interpretations.

## 7 | Participants

The design was a 2 (Consistency: inconsistent, consistent)  $\times$  2 (Evidence Type: report, video) with consistency as the

between-subjects factor and evidence type as the within-subjects factor.<sup>6</sup> An a priori power analysis using the *Superpower* package (Lakens and Caldwell 2020) in R (4.2.2; R Core Team 2022) indicated a sample size of 200 to detect a small to medium effect ( $f=0.15$ ),  $\alpha=0.05$ , and 96% power for the two-way interaction.

We recruited 349 United States Prolific workers. We excluded 50 due to a technical issue where participants could not view the officer's report, and 91 because they failed at least one manipulation check or did not complete the survey. The average time to complete the study was 31.29 min. The final sample consisted of 208 participants (56% male, 42% female, 3% non-binary) with an average age of 43.33 ( $SD=14.91$ ). Similarly to Experiment 1, our sample had higher education attainment (50%) compared to the average U.S. population (36%; McElrath 2025; see Table 1 for full demographics).

## 8 | Experiment 2 Results

We first assessed participants' views on pre-report review practices and replicated the findings of Experiment 1. Specifically, participants found it more reasonable for officers to review BWC footage *after* ( $M=5.73$ ) writing their reports than *before* ( $M=3.63$ ,  $t(206)=-7.51$ ,  $p<0.001$ , 95% CI =  $[-2.65, -1.55]$ ).

We use ratings for “How much do you trust that the *officer's statement described what actually happened?*” and “To what extent does the *officer's statement match the actual encounter?*” as the dependent variable for two analyses each. Though we did not pre-register it, we applied a Bonferroni corrected significance threshold of  $\alpha=0.025$  for the planned ANOVAs and regressions.

### 8.1 | Trust in Report and Footage

Did consistency influence trust in the officer's report? Yes. Across conditions, participants rated the BWC footage as more trustworthy than the officer's report, but this preference was stronger when the evidence was described as inconsistent ( $M_{\text{diff}}=2.15$ ,  $d=1.01$ ) than when described as consistent ( $M_{\text{diff}}=0.95$ ,  $d=0.50$ ). In other words, there was a significant evidence type and consistency interaction,  $F(1, 206)=18.38$ ,  $p<0.001$  (see Figure 3). In addition to the interaction, there was a main effect of evidence type,  $F(1, 206)=123.92$ ,  $p<0.001$ ,  $d=0.74$  ( $M_{\text{video}}=5.65$ ,  $M_{\text{report}}=4.10$ ). There was also a smaller and significant main effect of consistency ( $F(1, 206)=10.80$ ,  $p<0.001$ ,  $d=0.30$ ) such that evidence described as consistent ( $M=5.13$ ) was rated more trustworthy than evidence labeled as inconsistent ( $M=4.62$ ).

### 8.2 | Accuracy of Report and Footage

Our findings on accuracy were also replicated. Specifically, ratings of the extent to which the report or video matched the encounter depended on consistency ( $F(1, 206)=6.41$ ,  $p=0.012$ ; see Figure 4). Overall, participants rated the BWC footage higher, but in the inconsistent condition, the difference was greater ( $M_{\text{diff}}=1.30$ ,  $d=0.68$ ) than in the consistent condition ( $M_{\text{diff}}=0.63$ ,  $d=0.33$ ). There were main effects of consistency



( $F(1, 206)=7.00$ ,  $p=0.009$ ,  $d=0.27$ ) and evidence type ( $F(1, 206)=53.80$ ,  $p<0.001$ ,  $d=0.50$ ).

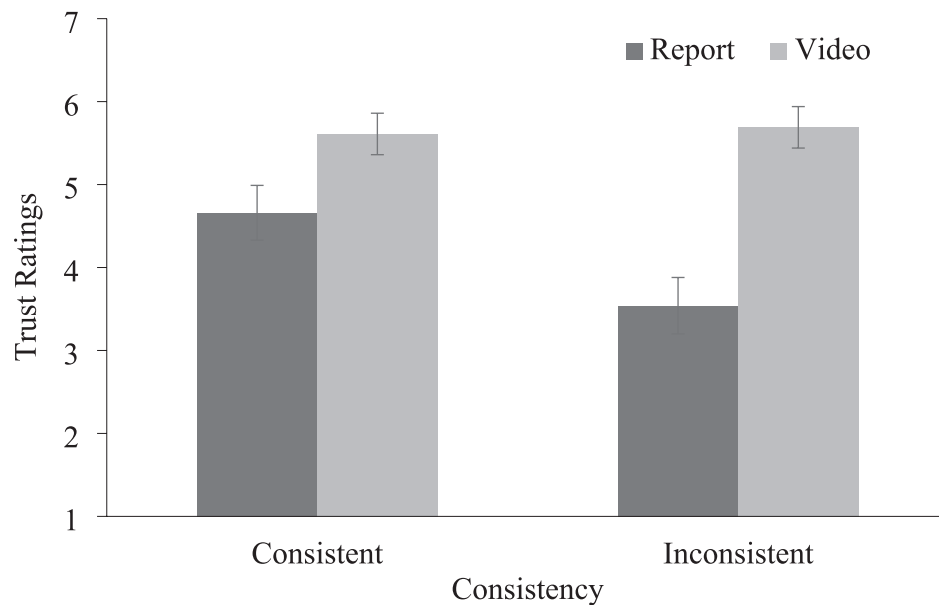
### 8.3 | IPS, Trust, and Accuracy

We tested the main effect of consistency on trust and accuracy. However, we wanted to examine whether IPS scores also predicted perceptions and interacted with consistency. A linear regression analysis for trust in the officer's report revealed that higher IPS scores were associated with greater trust in the officer's statement,  $b=0.49$ ,  $SE=0.10$ ,  $p<0.001$ . The interaction

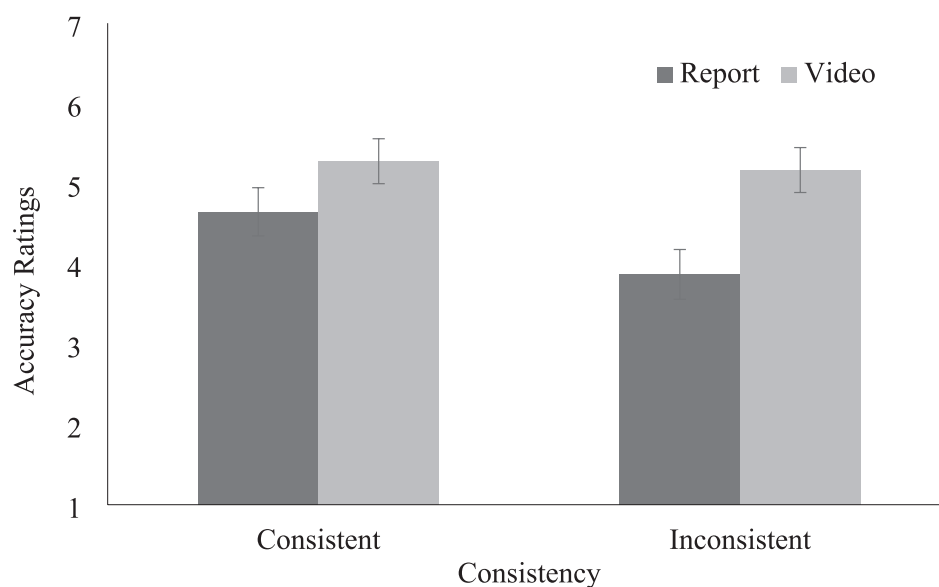
was not significant ( $p=0.50$ ). A second linear regression revealed that higher IPS scores were associated with greater accuracy ratings for the officer's report,  $b=0.38$ ,  $SE=0.10$ ,  $p<0.001$ . Here, too, the interaction was not significant ( $p=0.68$ ).

### 8.4 | Punishment Decisions and Officer Experience Ratings

Using regressions, we examined participant judgments about officer punishment with the same predictors as Experiment 1 (i.e., Review, Consistency, Review $\times$ Consistency, IPS, and



**FIGURE 3** | Experiment 2: Average perceived trustworthiness of the report and video based on consistency. Error bars represent 95% confidence intervals of cell means. The trust ratings were created as a within-subjects variable from the questions: “How much do you trust that *the officer's statement* described what actually happened?” and “How much do you trust that *the body camera footage* showed what actually happened?”



**FIGURE 4** | Experiment 2: Average beliefs about report and video accuracy based on consistency. Error bars represent 95% confidence intervals of cell means. The trust ratings were created as a within-subjects variable from the questions: “How much do you trust that *the officer's statement* described what actually happened?” and “How much do you trust that *the body camera footage* showed what actually happened?”

Consistency $\times$ IPS). When it comes to charging decisions, some of our results were replicated. The model was significant,  $F(3, 204) = 16.44$ ,  $p < 0.001$ . We again saw that participants with higher IPS were less likely to support charging the officer for their actions ( $b = -0.62$ ,  $SE = 0.12$ ,  $p > 0.001$ , 95% CI  $[-0.87, -0.39]$ ). However, there was no effect of consistency ( $p = 0.77$ ) or consistency $\times$ IPS ( $p = 0.79$ ; see Table S11).

Overall, 52.4% of participants did not think the officer should be found guilty if he was put on trial for his actions in the encounter. But did the predictors influence guilt decisions? The logistic regression revealed only a significant main effect of IPS scores,  $b = -0.52$ ,  $SE = 0.16$ ,  $p < 0.001$  (see Table S12). Specifically, as participants' IPS scores increased, the odds of a guilt judgment decreased by 40% (OR = 0.60, 95% CI  $[0.43, 0.79]$ ). There was no effect of consistency ( $p = 0.40$ ) or consistency $\times$ IPS ( $p = 0.72$ ).

Next, when participants were asked to choose which piece of evidence jurors should take into consideration most when making punishment decisions, 92.79% of participants picked the BWC footage over the officer's report. We then explored whether IPS scores influenced which evidence participants thought should take priority. A logistic regression with IPS as a predictor revealed the odds of prioritizing BWC footage decreased by 57% as IPS scores increased ( $\beta = -0.84$ ,  $SE = 0.27$ , OR = 0.42, 95% CI  $[0.25, 0.70]$ ,  $p = 0.002$ ).

We then examined participant rankings for reasons the officer said there was a knife, but no knife was visible. Like Experiment 1, most participants (46.08%) ranked "The officer thought he saw a knife" as the number one reason for the inconsistency. "The officer is lying about what he saw" was ranked number two for 40.20% of participants, while the technical issue explanation was ranked last by 47.06%. See Table S13 for a full breakdown of the rankings.

As an exploratory analysis, we conducted ordinal logistic regressions using IPS to further explore rankings of "The officer thought he saw a knife" and "The officer is lying about what he saw" explanations. Results indicated that participants who more strongly identified with police had greater odds of assigning the "The officer is lying" explanation as less plausible (i.e., a lower rank),  $\beta = 0.64$ ,  $SE = 0.16$ , OR = 1.90,  $p < 0.001$ , 95% CI  $[0.35, 0.97]$ . IPS did not significantly predict rank ratings for "The officer thought he saw a knife" ( $\beta = -0.24$ , OR = 0.79,  $p = 0.08$ ).

## 8.5 | Experiment 2 Summary

Overall, Experiment 2 largely replicated the findings from Experiment 1. First, consistency between the report and footage influenced participant ratings of trust and accuracy. However, unlike Experiment 1, consistency did not influence punishment decisions and perceptions of the officer's experience. But why did these findings not replicate? It is possible that the severity of use-of-force incidents played a role. In Experiment 1, the officer shot the civilian, whereas in Experiment 2, the officer beat the civilian with a baton. Perhaps using the baton was interpreted as less severe, as beatings could lead to less severe injuries, thus making the officer's actions less deserving of punishment. Although we did not assess the perceived severity of the officer's

actions, previous work suggests that people may view different types of force differently. In one study, participants rated baton use as less physically severe than gunfire in police-civilian encounters (Celestin and Kruschke 2019). However, Celestin and Kruschke (2019) used vignettes and did not assess punishment decisions. To our knowledge, no studies have directly tested whether the perceived severity of force influences officer punishment and accountability decisions.

Another possible answer for why consistency did not influence punishment decisions lies in the ambiguity of the footage. In Experiment 1, participants were told the report was inconsistent with the footage but never actually watched it themselves. Perhaps not seeing the footage meant participants imagined a clear inconsistency; an idealized version of what BWC footage could be. However, participants in Experiment 2 saw footage that did not confirm or deny the presence of a knife. Participants here may have assumed the knife was out of view, or they may have assumed the officer made a mistake. This interpretation is supported by our data, where we asked participants to rank explanations for the inconsistency. In both experiments, "the officer thought he saw a knife" was ranked as the number one reason for the inconsistency. Perhaps the ambiguity of the footage allowed participants to rationalize the inconsistency we mentioned as simply a mistake, thereby giving the officer the benefit of the doubt.

We also saw differences from Experiment 1 in how participants explained the inconsistencies in the officer's report. Specifically, there was increased skepticism for whether the officer was lying, which was the second highest explanation. In Experiment 1, this reason was ranked last. This data suggests a shift for participants who directly watched the BWC footage compared to when participants were simply provided with second-hand information.

Finally, IPS scores were associated with trust and accuracy ratings, charging decisions, and guilt judgments. Replicating Experiment 1, participants with higher IPS scores were less likely to advocate for officer punishment and wanted jurors to prioritize BWC footage less. Additionally, higher IPS scores were associated with higher ratings of trust and accuracy for the officer's report.

## 9 | General Discussion

BWCs have become increasingly central to how the media, public, and the legal system assess police behavior and accountability (e.g., Cullen 2018; Dimmick 2022; Yang 2019). Yet, what happens if what the officer says happened does not match the footage? Many officers worry that *any* inconsistencies between their reports and BWC footage will undermine their credibility. Given that BWCs are increasingly released to the media and used as evidence in courtrooms, understanding how discrepancies influence public perceptions and decisions is crucial.

Recall that across two experiments, we had participants read about non-fatal interactions where an officer claimed the civilian had a knife, prompting their use of force. We manipulated two key features: when the officer reviewed their BWC

footage (before or after their report) and whether that report was described as *consistent* or *inconsistent* with the footage. Participants then answered questions about trust, accuracy, and punishment. We aimed to answer several questions with our studies.

First, what were participants' perceptions of pre-report review, and did pre-report review influence perceptions of the officer? Overall, we found that participants showed a preference for post-report review. Specifically, they thought it was more reasonable for officers to write their reports *before* reviewing their BWC rather than after; a finding that mirrors researchers' and legal scholars' preferences (e.g., Farber 2024; Jones and Strange 2022). However, the review manipulation did not influence ratings of trust, accuracy, or punishment decisions. Researchers have asserted that pre-report review undermines the independence of evidence (Jones and Strange 2022; Pezdek 2015, 2022; Pezdek and Reisberg 2022). But departments encourage the practice *because* it will create consistency with the BWC footage. Our novel finding on people's perceptions of the pre-report policy suggests police are out on a limb; the public does not like pre-report review, but at the same time, they appear willing to defer to the fact it is a departmental policy.

Of course, experts know that a report that's written before viewing BWC footage is unlikely to be an exact duplicate of that BWC footage. The report and the footage are from different perspectives, and each has its own limitations. For example, BWCs do not capture everything the officer sees because they are usually mounted at chest height (Miller et al. 2014), and the report relies on human memory, which can be faulty due to inattentional blindness (Simons and Chabris 1999), susceptibility to misinformation (see Loftus 2005), and limitations on encoding. Despite these documented limitations, our results indicate people begin to lose trust in officers when their reports do not duplicate the footage. That result suggests people need may need to be educated on what is to be expected regarding inconsistencies.

Next, how did inconsistencies influence people's perceptions of the officer? As we predicted, in both experiments, when participants learned there was an inconsistency, they rated the officer's report as less trustworthy and accurate. One possible explanation for these results is what Dror (2020) calls the *fallacy of technological protection*—the mistaken belief that technology is inherently objective and unbiased. Supporting this idea, our participants provided significantly higher ratings of trust and accuracy for BWC footage. Perhaps people assume that BWCs capture the objective truth of the encounter, whereas the officer's report is subjective and therefore fallible. Indeed, we found that the highest ranked reason for the inconsistency was a mistake, that is “the officer thought he saw a knife.” Still, a significant minority of our participants—19% (Experiment 1) and 21% (Experiment 2)—ranked “the officer was lying” as the most likely explanation. Does the fallacy of technological protection contribute to diminished trust in the officer's report, especially when people are told there were inconsistencies? We suspect the answer is yes, but given we did not ask about people's beliefs regarding technology, it is an open question.

We also found that participants' Identification with Police (IPS) scores were significantly associated with all dependent

measures. In particular, higher IPS scores were associated with greater perceptions that the officer was accurate, trustworthy, and telling the truth (see Tables S4 and S9). Identifying with the police was also associated with less support for charging the officer and delivering a guilt judgment. Importantly, the associations we observed for IPS were not limited to perceptions; IPS was also associated with how people appear to fill in the gaps when information is ambiguous or missing. In Experiment 2, we inadvertently included a redundant question from Experiment 1: “The officer said that the civilian looked like he was going to lunge at him. How much do you believe that this is an accurate representation of what the officer experienced?” The report did not mention a lunge, and the footage did not show one. In fact, the civilian walks away from the officer. Yet, participants with higher IPS scores were more likely to endorse the idea that the officer perceived a threat (Table S10). Whether it was memory distortion due to our question phrasing (see Loftus 2005) or a reflection of a tendency to interpret missing information in favor of the officer remains an open question and certainly one worthy of future study.

## 10 | Limitations and Future Directions

We interpret our IPS findings cautiously because we measured IPS after our manipulations. It is possible that being told about inconsistencies influenced responses to the IPS. However, we do not think this is a likely explanation because our results follow the same patterns as previous studies where IPS influenced the interpretation of BWC footage (Granot et al. 2014; Jones, Crozier, and Strange 2017). If identification with police is indeed a pre-existing bias, then our results hint that there are downstream consequences for fairness, objectivity, interpretation of evidence, and guilt judgments, much like the cascading effects of a mistaken alibi or false confession (see Crozier et al. 2017; Kassin et al. 2025). Indeed, other studies show that those who strongly identify with police could show confirmation biases and selectively attend to evidence that is pro-police while discounting conflicting information (Oeberst and Imhoff 2023; Wason 1960). However, to confirm, future research should measure IPS before any manipulations or counterbalance when participants respond to the IPS to better assess how much IPS scores influence perceptions of trust, accuracy, and accountability.

A second limitation is the absence of a control group for the consistency manipulation due to our resource constraints (Lakens 2022). Without a control group, we are unable to determine whether simply mentioning an inconsistency would have affected participants' judgments. Although we chose to stick to the real-world scenario where if an officer's actions are being scrutinized, the consistency between the report and footage is likely to be raised by media or legal teams—the question remains, would participants have noticed the discrepancies themselves without our highlighting it? Would IPS scores have played a stronger role for a control group? Perhaps people with lower IPS scores would be more likely to identify inconsistencies, whereas people with higher IPS scores may overlook discrepancies and default to trusting the officer. Given the importance of answers to these questions for legal actors and policymakers, further investigation is warranted. If perceptions of an officer's credibility depend on whether inconsistencies are explicitly

identified, then expert testimony and juror education may need to address how to interpret inconsistencies and why they occur, a point we return to below.

Although consistency only predicted charging decisions and guilt judgments in Experiment 1, it uniformly affected trust and accuracy ratings. We caution against concluding that consistency has no effect on punishment decisions. In both studies, we told participants about a single inconsistency. We noted that the officer said there was a knife, but none was visible in the footage. In the real world, camera limitations, complexity of use-of-force incidents, and the reconstructive nature of memory are likely to produce multiple discrepancies. Thus, it is possible our studies underestimate the impact of discrepancies. We predict that as the number of inconsistencies increases, there will be more support for punishment and lower ratings for trust and accuracy. Future research should examine whether the amount or salience of inconsistencies impacts participant perceptions of officer trustworthiness, accuracy, and accountability.

Finally, when making their punishment decisions, we did not instruct participants to evaluate the officer's actions using the "objectively reasonable" standard (*Graham v. Connor* 1989), governing U.S. court decisions. This omission limits the generalizability of our findings. Would instructing participants to apply the "objectively reasonable" standard change their punishment decisions? That is an open question for future research, but we are inclined to predict it would not impact punishment judgments. Previous research suggests officers and civilians can view the same use-of-force incident and reach different conclusions about the justification of force (Pezdek et al. 2025). However, our findings do highlight what citizens might take away from information made available in the media, which has direct implications for community and police relations.

In summary, our data show that consistency matters, even though consistency is unlikely. Although our findings show inconsistencies can reduce public trust, this should not be interpreted as support for pre-report review policies. Writing the report after reviewing the footage will reduce discrepancies but can introduce a new set of problems. Prior research shows that pre-report review influences what is included in the report (Adams et al. 2020). In other words, when consistency is created through pre-report review, it undermines the independence of the evidence.

Some agencies have begun using solutions like automated report writing using Artificial Intelligence (AI; Booth 2024; Santa Cruz 2025). The tool generates narrative reports based on the audio of the BWC footage, promising faster reports (Axon n.d.) and potentially reducing discrepancies. But these practices also raise concerns. First, they rely on a limited perspective; BWC audio may omit important contextual information. Second, they may exacerbate the fallacy of technological protection (Dror 2020). Third, the only independent, peer-reviewed testing to our knowledge has focused on report writing speed (Adams et al. 2024) and found there were no time savings. In short, it remains unclear how these AI reports affect public perceptions of credibility, trust, and whether there are consequences for officer memory. Automating reports may simply offer an illusion of accuracy and neutrality.

Ultimately, we believe policies should aim to preserve the independence of what the officer says happened and what the camera records. Instead of focusing on eliminating inconsistencies, we should focus on helping the public and legal actors understand why they occur, how to interpret them, and when they warrant concerns. Future research should examine potential interventions to mitigate the effect of consistency on trust, accuracy, and punishment decisions when the video is unavailable. Would interventions such as jury instructions, expert testimony, or educational materials help educate the public on the limitations of BWC footage? We know from prior research that judicial safeguards can help the general public understand the fallibility of eyewitness memory (Cutler et al. 1989; Jones, Bergold, et al. 2017). If current safeguards like jury instructions and expert testimony could sensitize the general public to the limitations of BWC footage, then police could modify their current policies regarding pre-report review. Doing so just might preserve the independence of their evidence (*Graham v. Connor* 1989) and bring back the promise that BWC footage can promote transparency and trust. But if current safeguards do not work, then we will need to investigate new approaches to mitigate the damage.

## 11 | Conclusion

Body-worn cameras are not going away. Instead, their use has extended to correctional facilities (e.g., Lawrence et al. 2024), immigration enforcement (U.S. Immigration and Customs Enforcement 2024), the Department of Homeland Security (Department of Homeland Security 2023), health-care (Spears 2024), and eldercare facilities (Span 2025). Internationally, bus drivers, paramedics (The Economist 2018), and campus security officers (Menichelli et al. 2024) also wear BWCs. As BWCs spread, the consequences of potential discrepancies in evidence will become more widespread. Our findings suggest that a single discrepancy can begin to erode trust. Without the public understanding the limitations of each evidence source, they may continue to overvalue consistency, reducing credibility for what are likely explainable discrepancies.

### Author Contributions

**Kris-Ann S. Anderson:** writing – original draft, formal analysis, investigation, visualization, validation, data curation, writing – review and editing. **Kristyn Jones:** formal analysis, investigation, conceptualization, methodology, data curation, project administration. **Deryn Strange:** writing – review and editing, conceptualization, methodology, funding acquisition.

### Ethics Statement

Both experiments were approved by the John Jay College of Criminal Justice's Human Research Protection Program.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Data Availability Statement

The pre-registration (<https://osf.io/bygx2>), data, and materials are available on the Open Science Framework at <https://osf.io/j9d37/>.



## Endnotes

- <sup>1</sup>In the United States, *Graham v. Connor* (1989) is the U.S. Supreme Court decision that established the legal standard to evaluate police use-of-force incidents.
- <sup>2</sup>We used features to block duplicate IP addresses, block suspicious geocode locations, and verify worker country location to improve data quality.
- <sup>3</sup>These two accuracy questions had high positive correlations with other measures ( $r^2=0.77-0.90$ ), we have included these analyses in Table S3.
- <sup>4</sup>This question on beliefs about whether the officer was telling the truth was highly correlated with the trust and accuracy measures ( $r^2=0.83-0.90$ ). As such, these analyses are in Tables S4 and S9.
- <sup>5</sup>A technical glitch meant some participants did not receive this question.
- <sup>6</sup>We excluded BWC Review from the ANOVA analyses in Experiment 2 based on the sensitivity analysis and resource constraints (Lakens 2022). We did however include BWC review in a series of exploratory analysis to determine if those findings were replicated (see Tables S6–S8).

## References

- Adams, D., H. M. Paterson, and H. G. MacDougall. 2020. "Law and (Re)order: Updating Memory for Criminal Events With Body-Worn Cameras." *PLoS One* 15, no. 12: e0243226. <https://doi.org/10.1371/journal.pone.0243226>.
- Adams, I. T., M. Barter, K. McLean, H. M. Boehme, and I. A. Geary. 2024. "No Man's Hand: Artificial Intelligence Does Not Improve Police Report Writing Speed." *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-024-09644-7>.
- Akehurst, L., G. Köhnken, A. Vrij, and R. Bull. 1996. "Lay Persons' and Police Officers' Beliefs Regarding Deceptive Behaviour." *Applied Cognitive Psychology* 10, no. 6: 461–471. [https://doi.org/10.1002/\(SICI\)1099-0720\(199612\)10:6<461::AID-ACP413>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-0720(199612)10:6<461::AID-ACP413>3.0.CO;2-2).
- Axon. n.d. "Draft One." Axon. <https://www.axon.com/products/draft-one>.
- Benner, K. 2019. "Eric Garner's Death Will Not Lead to Federal Charges for N.Y.P.D Officer." *The New York Times*. <https://www.nytimes.com/2019/07/16/nyregion/eric-garner-daniel-pantaleo.html>.
- Berman, G. L., D. J. Narby, and B. L. Cutler. 1995. "Effects of Inconsistent Eyewitness Statements on Mock-Jurors' Evaluations of the Eyewitness, Perceptions of Defendant Culpability and Verdicts." *Law and Human Behavior* 19, no. 1: 79–88. <https://doi.org/10.1007/BF01499074>.
- Booth, B. 2024. "Police Departments Across U.S. Are Starting to Use Artificial Intelligence to Write Crime Reports." *CNBC*. <https://www.cnbc.com/2024/11/26/police-departments-are-using-ai-to-write-crime-reports.html>.
- Brewer, N., R. Potter, R. P. Fisher, N. Bond, and M. A. Luszcz. 1999. "Beliefs and Data on the Relationship Between Consistency and Accuracy of Eyewitness Testimony." *Applied Cognitive Psychology* 13, no. 4: 297–313. [https://doi.org/10.1002/\(SICI\)1099-0720\(199908\)13:4<297::AID-ACP578>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<297::AID-ACP578>3.0.CO;2-S).
- Brysbaert, M. 2019. "How Many Participants Do We Have to Include in Properly Powered Experiments? A Tutorial of Power Analysis With Reference Tables." *Journal of Cognition* 2, no. 1: 16. <https://doi.org/10.5334/joc.72>.
- Celestin, B. D., and J. K. Kruschke. 2019. "Lay Evaluations of Police and Civilian Use of Force: Action Severity Scales." *Law and Human Behavior* 43, no. 3: 290–305. <https://doi.org/10.1037/lhb0000333>.
- Chapman, B. 2019. "Body-Worn Cameras: What the Evidence Tells Us." *Nij Journal*. 280. <https://nij.ojp.gov/topics/articles/body-worn-cameras-what-evidence-tells-us>.
- Crozier, W. E., D. Strange, and E. F. Loftus. 2017. "Memory Errors in Alibi Generation: How an Alibi Can Turn Against Us." *Behavioral Sciences & the Law* 35, no. 1: 6–17. <https://doi.org/10.1002/bsl.2273>.
- Cullen, T. 2018. "Alabama Cop Cleared of Shooting Man Brandishing Wallet as Dashcam Video Is Released." *Daily News*. <https://www.nydailynews.com/news/national/cleared-shooting-man-waving-wallet-graphic-video-article-1.2995182>.
- Cutler, B. L., H. R. Dexter, and S. D. Penrod. 1989. "Expert Testimony and Jury Decision Making: An Empirical Analysis." *Behavioral Sciences & the Law* 7, no. 2: 215–225. <https://doi.org/10.1002/bsl.2370070206>.
- Deeb, H., A. Vrij, L. Hope, S. Mann, P. A. Granhag, and L. A. Strömwall. 2018. "Police Officers' Perceptions of Statement Inconsistency." *Criminal Justice and Behavior* 45, no. 5: 644–665. <https://doi.org/10.1177/0093854818758808>.
- Department of Homeland Security. 2023. "DHS Announces First Department-Wide Policy on Body-Worn Cameras. U.S." Department of Homeland Security. <https://www.dhs.gov/archive/news/2023/05/23/dhs-announces-first-department-wide-policy-body-worn-cameras>.
- Dimmick, I. 2022. "SAPD Fires Its Rookie Officer Who Shot, Injured 17-Year-Old. San Antonio Report." <https://sanantonioreport.org/sapd-rookie-officer-termination-shooting-teenager/>.
- Dror, I. E. 2020. "Cognitive and Human Factors in Expert Decision Making: Six Fallacies and the Eight Sources of Bias." *Analytical Chemistry* 92, no. 12: 7998–8004. <https://doi.org/10.1021/acs.analchem.0c00704>.
- Estate of Miguel Antonio Richards v. City of New York. 2024. "No. 1:18-cv-11287." <https://law.justia.com/cases/federal/district-courts/new-york/nysdce/1:2018cv11287/505862/113/>.
- Farber, H. B. 2024. "Write Before You Watch: Policies for Police Body-Worn Cameras That Advance Accountability and Accuracy." *American Criminal Law Review* 61, no. 1: 59–94. <https://doi.org/10.2139/ssrn.4778640>.
- Farooq, U. 2023. "Body Cameras Were Sold as a Tool of Police Reform. Ten Years Later, Most of the Footage is Kept From Public View. ProPublica." <https://www.propublica.org/article/body-camera-videos-police-killings-remain-hidden-from-public>.
- Faul, F., E. Erdfelder, A. G. Lang, and A. Buchner. 2007. "G\*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39, no. 2: 175–191. <https://doi.org/10.3758/bf03193146>.
- Fisher, R. P., N. Brewer, and G. Mitchell. 2009. "The Relation Between Consistency and Accuracy of Eyewitness Testimony: Legal Versus Cognitive Explanations." In *Handbook of Psychology of Investigative Interviewing: Current Developments and Future Directions*, edited by R. Bull, T. Valentine, and T. Williamson, 121–136. Wiley Blackwell. <https://doi.org/10.1002/9780470747599.ch8>.
- Grady, R. H., B. J. Butler, and E. F. Loftus. 2016. "What Should Happen After an Officer-Involved Shooting? Memory Concerns in Police Reporting Procedures." *Journal of Applied Research in Memory and Cognition* 5, no. 3: 246–251. <https://doi.org/10.1016/j.jarmac.2016.07.002>.
- Graham v. Connor. 1989. "490 U.S. 386."
- Granhag, P. A., and L. A. Strömwall. 2000. "Effects of Preconceptions on Deception Detection and New Answers to Why Lie-Catchers Often Fail." *Psychology, Crime & Law* 6, no. 3: 197–218. <https://doi.org/10.1080/10683160008409804>.
- Granot, Y., E. Balcetis, K. E. Schneider, and T. R. Tyler. 2014. "Justice Is Not Blind: Visual Attention Exaggerates Effects of Group Identification on Legal Punishment." *Journal of Experimental Psychology: General* 143, no. 6: 2196–2208. <https://doi.org/10.1037/a0037893>.

- Hermann, P., and R. Weiner. 2014. "Issues Over Police Shooting in Ferguson Lead Push for Officers and Body Cameras. Washington Post." [https://www.washingtonpost.com/local/crime/issues-over-police-shooting-in-ferguson-lead-push-for-officers-and-body%20cameras/2014/12/02/dedcb2d8-7a58-11e4-84d4-7c896b90abdc\\_story.html](https://www.washingtonpost.com/local/crime/issues-over-police-shooting-in-ferguson-lead-push-for-officers-and-body%20cameras/2014/12/02/dedcb2d8-7a58-11e4-84d4-7c896b90abdc_story.html).
- House, W. 2022. "Fact Sheet: President Biden's Safer America Plan. The White House. [Press Release]." <https://web.archive.org/web/20220801192916/https://www.whitehouse.gov/briefing-room/statements-releases/2022/08/01/fact-sheet-president-bidens-safer-america-plan-2/>.
- Hutto, J. W., and R. D. Green. 2016. "Social Movements Against Racist Police Brutality and Department of Justice Intervention in Prince George's County, Maryland." *Journal of Urban Health* 93, no. 1: 89. <https://doi.org/10.1007/s11524-015-0013-x>.
- Hyland, S. S. 2018. "Body-Worn Cameras in Law Enforcement Agencies, 2016 (NCJ 251775)." Bureau of Justice Statistics, U.S. Department of Justice. <https://bjs.ojp.gov/content/pub/pdf/bwclea16.pdf>.
- Jones, A. M., A. N. Bergold, M. K. Dillon, and S. D. Penrod. 2017. "Comparing the Effectiveness of Henderson Instructions and Expert Testimony: Which Safeguard Improves Jurors' Evaluations of Eyewitness Evidence?" *Journal of Experimental Criminology* 13, no. 1: 29–52. <https://doi.org/10.1007/s11292-016-9279-6>.
- Jones, K. A., W. E. Crozier, and D. Strange. 2017. "Believing Is Seeing: Biased Viewing of Body-Worn Camera Footage." *Journal of Applied Research in Memory and Cognition* 6, no. 4: 460–474. <https://doi.org/10.1016/j.jarmac.2017.07.007>.
- Jones, K. A., and D. Strange. 2022. "Officer Memory Could Be Tainted by BWC Footage; So, What Is the Solution?" *Journal of Applied Research in Memory and Cognition* 11, no. 2: 166–170. <https://doi.org/10.1037/mac0000040>.
- Kassin, S. M., H. M. D. Cleary, G. H. Gudjonsson, et al. 2025. "Police-Induced Confessions, 2.0: Risk Factors and Recommendations." *Law and Human Behavior* 49, no. 1: 7–53. <https://doi.org/10.1037/lhb0000593>.
- Lakens, D. 2022. "Sample Size Justification." *Collabra: Psychology* 8, no. 1: 33267. <https://doi.org/10.1525/collabra.33267>.
- Lakens, D., and A. Caldwell. 2020. "Superpower: Simulation-Based Power Analysis for Factorial Designs (R package Version 0.1.0) [Computer software]." <https://cran.r-project.org/web/packages/Superpower/index.html>.
- Lawrence, D. S., B. E. Peterson, M. D. White, B. C. Cunningham, and J. R. Coldren Jr. 2024. "Effects of Correctional Body-Worn Cameras on Responses to Resistance: A Randomized Controlled Trial in a Jail Setting." *Justice Quarterly* 41, no. 6: 892–914. <https://doi.org/10.1080/07418825.2024.2327301>.
- Litman, L., J. Robinson, and T. Abberbock. 2017. "TurkPrime.com: A Versatile Crowdsourcing Data Acquisition Platform for the Behavioral Sciences." *Behavior Research Methods* 49, no. 2: 433–442. <https://doi.org/10.3758/s13428-016-0727-z>.
- Loftus, E. F. 2005. "Planting Misinformation in the Human Mind: A 30-Year Investigation of the Malleability of Memory." *Learning & Memory* 12, no. 4: 361–366. <https://doi.org/10.1101/lm.94705>.
- Lopez, G. 2023. "A Disappointing Policing Change. Police-Worn Body Cameras Have Not Lived Up to Expectations. The New York Times." <https://www.nytimes.com/2023/12/14/briefing/police-body-cameras.html>.
- Lowenkron, H. 2021. "A Wellness Check Killed Miguel Richards. His Family Still Questions Why. The New York Times Student Journalism Institute. Medium." <https://nytimes-institute.medium.com/miguel-richards-8345b560b100>.
- Lum, C., C. S. Koper, D. B. Wilson, et al. 2020. "Body-Worn Cameras' Effects on Police Officers and Citizen Behavior: A Systematic Review." *Campbell Systematic Reviews* 16, no. 3: e1112. <https://doi.org/10.1002/cl2.1112>.
- McElrath, K. 2025. "Field of Bachelor's Degree in the United States: 2022. United States Census Bureau." <https://www.census.gov/library/publications/2025/acs/acs-59.html>.
- Menichelli, F., K. Bullock, J. Garland, and J. Allen. 2024. "Policing Universities: Exploring the Use of Body-Worn Cameras (BWCs) by Private Campus Security Officers." *Policing and Society* 34, no. 7: 658–673. <https://doi.org/10.1080/10439463.2024.2315583>.
- Miller, L., J. Toliver, and P. E. R. Forum. 2014. *Implementing a Body-Worn Camera Program: Recommendations and Lessons Learned*. U.S. Department of Justice, Bureau of Justice Assistance, Office of Community Oriented Policing Services. [https://www.policeforum.org/assets/docs/Free\\_Online\\_Documents/Technology/implementing%20a%20body-worn%20camera%20program.pdf](https://www.policeforum.org/assets/docs/Free_Online_Documents/Technology/implementing%20a%20body-worn%20camera%20program.pdf).
- Mustian, J. 2018. "It's Unfortunate: Defective Body Cam Didn't Record Trooper Shooting Baton Rouge Teen. The Advocate." [https://www.theadvocate.com/baton-rouge/news/crime-police/article\\_e0d3f984-9106-11e8-914a-5766f188bfa3.html](https://www.theadvocate.com/baton-rouge/news/crime-police/article_e0d3f984-9106-11e8-914a-5766f188bfa3.html).
- New York City Police Department. 2017. "NYPD Releases Body-Worn Camera Footage Captured During a Police-Involved Shooting on September 6, 2017." <https://www.nyc.gov/site/nypd/news/article.page?id=pr0914&permalinkName=nypd-releases-body-worn-camera-footage-captured-during-police-involved-shooting-september-6>.
- Oeberst, A., and R. Imhoff. 2023. "Toward Parsimony in Bias Research: A Proposed Common Framework of Belief-Consistent Information Processing for a Set of Biases." *Perspectives on Psychological Science* 18, no. 6: 1464–1487. <https://doi.org/10.1177/17456916221148147>.
- Office of Justice Programs. 2016. "Body-Worn Cameras in Law Enforcement Agencies. U.S. Department of Justice." <https://www.bjs.gov/content/pub/pdf/bwclea16.pdf>.
- Palmer, M. A., L. Button, E. Barnett, and N. Brewer. 2016. "Inconsistencies Undermine the Credibility of Confession Evidence." *Legal and Criminological Psychology* 21, no. 1: 161–173. <https://doi.org/10.1111/lcrp.12048>.
- Pezdek, K. 2015. "Should Cops Get to Review the Video Before They Report? The Marshall Project." <https://www.themarshallproject.org/2015/08/13/should-cops-get-to-review-the-video-before-they-report>.
- Pezdek, K. 2022. "Psychological Research on the Use of Body-Worn Cameras." In *Advances in Psychology and Law*, edited by B. H. Bornstein, M. K. Miller, and D. DeMatteo, vol. 6, 39–62. Springer.
- Pezdek, K., and D. Reisberg. 2022. "Psychological Myths About Evidence in the Legal System: How Should Researchers Respond?" *Journal of Applied Research in Memory and Cognition* 11, no. 2: 143–156. <https://doi.org/10.1037/mac0000037>.
- Pezdek, K., T. Shapland, and J. Barragan. 2025. "Perceptions of Officer-Involved Shootings by Police Officers Versus Civilians." *Psychology, Crime & Law* 31, no. 8: 938–961. <https://doi.org/10.1080/1068316X.2024.2319313>.
- R Core Team. 2022. "R: A Language and Environment for Statistical Computing (Version 4.2.2) [Computer Software]." R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Santa Cruz, E. 2025. "Arizona Police Testing Axon AI Software to Draft Reports From Bodycam Footage. AZ Central." <https://www.azcentral.com/story/news/local/arizona/2025/04/01/arizona-police-testing-ai-to-draft-reports-from-bodycam-footage/79343735007/>.
- Simons, D. J., and C. F. Chabris. 1999. "Gorillas in Our Midst: Sustained Inattention Blindness for Dynamic Events." *Perception* 28, no. 9: 1059–1074. <https://doi.org/10.1068/p281059>.
- Span, P. 2025. "Why Cameras Are Popping up in Eldercare Facilities. The New York Times." <https://www.nytimes.com/2025/04/07/health/cameras-assisted-living.html>.

Spears, S. J. 2024. "Body-Worn Cameras in Healthcare (IAHSS-F RS-24-01). International Association for Healthcare Security and Safety (IAHSS) Foundation." <https://iahssf.org/research/body-worn-cameras-in-healthcare-policies-regarding-homeless-populations-in-hospitals-cloned/>.

Stinson, P. M., and C. Wentzlof. 2019. *Research Brief One-Sheet—No. 9: On-Duty Shootings: Police Officers Charged With Murder or Manslaughter, 2005-2019*. Criminal Justice Faculty Publications. [https://scholarworks.bgsu.edu/cgi/viewcontent.cgi?article=1100&context=crim\\_just\\_pub](https://scholarworks.bgsu.edu/cgi/viewcontent.cgi?article=1100&context=crim_just_pub).

Tavernise, S. (Host). 2024. *The Failed Promise of Police Body Cameras* [Audio Podcast]. The New York Times.

The Economist. 2018. "Body-Worn Cameras Are Spreading Beyond the Police. The Economist." <https://www.economist.com/britain/2018/07/28/body-worn-cameras-are-spreading-beyond-the-police>.

Todak, N., L. Leban, and L. James. 2021. "Citizen Attitudes Towards the Public Release of Police Body-Worn Camera Video Footage." *Police Practice and Research* 22, no. 7: 1760–1776. <https://doi.org/10.1080/15614263.2021.1969929>.

Tyler, T. R., and J. Fagan. 2008. "Legitimacy and Coooperation: Why Do People Help the Police Fight Crime in Their Communities?" *Ohio State Journal of Criminal Law* 6: 231–275. [https://scholarship.law.columbia.edu/faculty\\_scholarship/414](https://scholarship.law.columbia.edu/faculty_scholarship/414).

Umansky, E. 2023. "The Failed Promise of Police Body Cameras. The New York Times." <https://www.nytimes.com/2023/12/13/magazine/police-body-cameras-miguel-richards.html>.

Upturn. 2017. "The Illusion of Accuracy: How Body-Worn Camera Footage Can Distort Evidence." <https://www.upturn.org/reports/2017/the-illusion-of-accuracy/>.

U.S. Immigration and Customs Enforcement. 2024. "ICE Announces Initial Deployment of Body-Worn Cameras." <https://www.ice.gov/news/releases/ice-announces-initial-deployment-body-worn-cameras>.

Vredeveldt, A., L. Kesteloo, and A. Hildebrandt. 2021. "To Watch or Not to Watch: When Reviewing Body-Worn Camera Footage Improves Police Reports." *Law and Human Behavior* 45, no. 5: 427–439. <https://doi.org/10.1037/lhb0000462>.

Wason, P. C. 1960. "On the Failure to Eliminate Hypotheses in a Conceptual Task." *Quarterly Journal of Experimental Psychology* 12: 129–140. <https://doi.org/10.1080/17470216008416717>.

White House. 2024. "Budget of the U.S. Government: Fiscal Year 2023." [https://www.whitehouse.gov/wp-content/uploads/2022/03/budget\\_fy2023.pdf](https://www.whitehouse.gov/wp-content/uploads/2022/03/budget_fy2023.pdf).

Wright, J. E., and A. M. Headley. 2021. "Can Technology Work for Policing? Citizen Perceptions of Police-Body Worn Cameras." *American Review of Public Administration* 51, no. 1: 17–27. <https://doi.org/10.1177/0275074020945632>.

Yang, A. 2019. "Officers Cleared of Wrongdoing in Shooting That Left Firefighter, Suspect Dead in Wisconsin. ABC News." <https://abcnews.go.com/US/officers-cleared-wrongdoing-shooting-left-firefighter-suspect-dead/story?id=63717912>.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** acp70113-sup-0001-Supinfo.docx.