Force and Fallout: Experimental Evidence from a National Test of the Community Expectations Standard

Christopher J. Marier[a,*], Conor Goodwin[b]

[a] *University of Central Florida, 12805 Pegasus Dr., Orlando, 32816, FL, United States*
[b] *Appalachian State University, 224 Joyce Lawrence Lane, Boone, 28608, NC, United States*

Abstract

A persistent "reasonableness divide" exists between the legal standards governing police use of force and the public's expectations, producing "lawful but awful" uses of force. This study empirically tests the "Community Expectations Standard" (CES), a model that identifies five criteria the public uses to evaluate force: underlying governmental interest, avoidability, officer motivation, subject resistance, and the presence of a "highly dangerous" environment. Using a factorial survey experiment with a national sample of nearly 2,000 U.S. adults, we analyzed responses to a hypothetical vignette depicting non-lethal force using t-tests and Bayesian linear regression models. The results show that underlying governmental interest and subject resistance—those factors with analogs in constitutional law—are the most powerful predictors of reasonableness judgments. In contrast, other CES factors were weaker, with their effects potentially filtered through the observer's personal characteristics. Notably, political partisanship emerged as a more potent predictor of reasonableness appraisals than race or ethnicity, suggesting partisanship acts as a primary lens for interpreting police use of force. We conclude that the CES framework is a valuable tool but should be refined to distinguish between objective, event-based criteria and subjective, observer-based criteria. Bridging the reasonableness divide requires adjusting both law enforcement practices and public expectations.

Keywords: use of force, Community Expectations Standard, Reasonableness divide, Public perception, Police legitimacy, Political partisanship

- The final publisher's manuscript can be downloaded from the journal website: https://doi.org/10.1016/j.jcrimjus.2025.102571
- This research was supported by a grant from the University Research Council at Appalachian State University

*Corresponding author
Email address: christopher.marier@ucf.edu (Christopher J. Marier)

A fundamental tension exists in most developed societies: the use or threat of physical coercion is simultaneously necessary and reviled. Coercive authority is perhaps the defining feature of police work: in contemporary societies, private citizens have largely given up their right to use violence to resolve conflicts, and instead concentrated it within an institution with explicit, bounded authority to use or threaten physical coercion, when necessary, to compel people to comport with pre-established rules of conduct (Bittner, 1990; Klockars, 1985). While this has undoubtedly reduced the sort of violence and vigilantism that has historically characterized social conflict resolution (Pinker, 2012), the institutionalization of coercive authority is problematic when those institutions lose legitimacy in the eyes of the public (Levitsky and Ziblatt, 2019). That legitimacy is largely tied to perceptions of institutional accountability and erodes when citizens do not believe institutions are holding themselves to acceptable standards (Peters and Pierre, 2006).

The coercive authority of the police is not absolute, and police use of force is evaluated against multiple criteria, including agency policy, state law, federal statutes, and constitutional case law (Stoughton et al., 2021). However, police violence is also evaluated in the court of public opinion. Public trust in government institutions, including the police, remains near historic lows (Saad, 2023). Nearly two thirds of Americans believe that police violence is a serious problem (Smith, 2020). Yet, between 2005 and 2019 just 35 officers have been convicted of a crime related to an on-duty shooting (Stinson and Wentzlof, 2019), despite shooting and killing about 1,000 people per year.[1] Thus, a majority of Americans think that the police use of force is often problematic, yet legal outcomes rarely converge with this sentiment. The phrase "lawful but awful" is sometimes used to describe situations where a police officer's use of force is legally justified under existing laws and standards, yet it appears excessive, unnecessary, or morally questionable to the public. These actions, while technically within the boundaries of the law, are often perceived as troubling or unacceptable by community members, producing what has been called a "reasonableness divide" between legal institutions and the public (Richardson and Fridell, 2025; Fridell et al., 2023). This has led to demands for dramatic policy reforms, some going so far as to advocate for defunding or abolishing the police (Vitale, 2021; Maher, 2021). While courts and law enforcement professionals emphasize adherence to legal standards, community members set informal, normative standards of conduct through which the police are evaluated.

In short, even legally reasonable police use of force is commonly viewed as unjustified by the public (Mourtgos and Adams, 2020; Richardson and Fridell, 2025). So just how do community members evaluate police use of force, if not with reference to the law? Only recently has a model been devised to help explain the criteria through which the public commonly evaluates police force: the "Community Expectations Standard" (CES), which illustrates tensions with legal and administrative standards (Stoughton et al., 2021). Yet little systematic

[1] https://github.com/washingtonpost/data-police-shootings

examination has tested these Community Expectation Standards (cf. Richardson and Fridell 2025; Fridell and Marier 2023), and none have experimentally evaluated all five domains proposed by the model. The CES model may help explain the association of personal attributes with perceptions of police conduct and systematically explain tension between the law and public opinion.

This study empirically evaluates the Community Expectations Standard in a factorial survey experiment using a national sample of approximately 2,000 American adults.

1. Review

For decades, research has documented *that* demographic and political identities predict public opinion on police force, but has spent less attention to explaining *why*. Women, for example, consistently express lower levels of support for police force compared to men (e.g., Thompson and Lee 2004; Trahan and Russell 2017). Political ideology also shapes evaluations: conservatives are generally more supportive of police use of force than are liberals or moderates (e.g., Braga et al. 2014; Carter and Corra 2016; Carter et al. 2016). Similarly, individuals who report greater trust in, identification with, or perceptions of legitimacy toward police are more likely to view both actual and hypothetical uses of force as justified (e.g., Bradford et al. 2017; Gerber and Jackson 2017; cf. Baker and Bacharach 2017; Choi 2019). In contrast, findings related to age, education, and income have been mixed (see Richardson 2022, Chapter 2).

Race is among the most consistent predictors of community evaluations of police force. Across diverse contexts, White Americans tend to support police use of force at significantly higher rates than Black Americans (Mourtgos and Adams, 2020; Silver and Pickett, 2015; Simon et al., 2021; Strickler and Lawson, 2022). In an experimental study, for instance, participants' demographic characteristics—especially race—were the primary drivers of their assessments of excessive force, outweighing factors such as officer gender or contextual factors (Baker and Bacharach, 2017). Fridell and Marier (2023) similarly found that Black respondents were less likely than Whites to judge an officer's use of force as reasonable, particularly when the subject shot was Black. In another experiment, White respondents rated a police shooting as more justifiable than other respondents (Marier, 2024). Analysis of more than three decades of General Social Survey data shows that Black Americans are about half as likely as Whites to agree that it is *ever* appropriate for police to strike a citizen, even in legally justified scenarios such as resisting arrest or attacking an officer (Mourtgos and Adams, 2020; Silver and Pickett, 2015).

While a substantial body of research has thus documented the ways personal characteristics and attitudes intersect with perceptions of the police, there has been little effort to systematically explain the criteria that produce these divergent evaluations. Most studies describe *that* these differences exist without providing a theoretical model to make sense of *why* they exist. The Community Expectations Standard fills this theoretical gap by identifying criteria through which the public commonly evaluates police force and illustrating the

tensions with legal and administrative standards. The CES model moves beyond associating personal attributes with perceptions and instead attempts to systematically explain the tension between law and public opinion, providing a framework for understanding why legally justified actions are so often met with public condemnation.

The "reasonableness divide" between public perception and legal outcomes stems from the application of different evaluative frameworks (Stoughton et al. 2021; Richardson and Fridell 2025; cf. Mourtgos et al. 2025). While police actions are formally judged against a complex web of legal and administrative rules, the public often applies an informal, normative standard rooted in societal values and expectations.

To systematically analyze the tension that the Community Expectations Standard model seeks to explain, it is first necessary to articulate each of these distinct evaluative frameworks. The following sections therefore detail the primary standards against which police use of force is judged: the legal standards derived from constitutional and statutory law, and the Community Expectations Standard that captures public perception.

1.1. Legal Standards

Police use of force is evaluated under constitutional law primarily through the lens of the Fourth Amendment, which protects individuals against "unreasonable seizures" by the government. The prevailing constitutional standard was articulated in *Graham v. Connor* (1989), where the U.S. Supreme Court held that claims of excessive force by law enforcement during an arrest or investigatory stop should be analyzed under the Fourth Amendment's "objective reasonableness" standard. This standard assesses whether an officer's actions are reasonable in light of the facts and circumstances confronting them, without regard to their underlying intent or motivation (*Graham v. Connor*, 490 U.S. 386, 1989).

The *Graham* decision identified three core factors relevant to determining reasonableness: the severity of the crime at issue; whether the suspect posed an immediate threat to the safety of the officers or others; and whether the suspect was actively resisting arrest or attempting to evade arrest by flight. Importantly, the Court emphasized that the reasonableness inquiry must be judged from the perspective of a reasonable officer on the scene, rather than with the benefit of hindsight, and that officers are often forced to make split-second decisions under tense and rapidly evolving circumstances.

"Objective reasonableness" must consider proportionality and the totality of the circumstances, including factors such as the availability of less intrusive means and the foreseeability of harm. In *Tennessee v. Garner* (1985), the Court held that deadly force may not be used against a fleeing suspect unless the officer has probable cause to believe that the suspect poses a significant threat of death or serious physical injury to the officer or others (*Tennessee v. Garner*, 471 U.S. 1, 1985). The U.S. Supreme Court's recent decision in *Barnes v. Felix* (2025) rejected the "moment of threat" doctrine, which only considers the circumstances at the precise moment an officer perceives danger. Instead,

the Court affirmed the need to evaluate the use of force under the "totality of the circumstances." This means courts must consider all relevant factors, including what led up to the incident, not just the immediate moment of the use of force.

Officers are not judged by what they actually believed, but rather by what a hypothetical reasonable officer in the same situation would have perceived and done. This framework permits force that is mistaken but reasonable, and it rejects the use of force that is excessive in light of the government's interests: enforcing the law, maintaining order, and ensuring officer safety. Courts generally defer to officers' on-the-ground judgments unless the force used is clearly excessive or unjustified under prevailing standards.

Under federal law 42 U.S.C. § 1983, citizens may hold a police officer civilly liable for a constitutional violation in either state or federal courts. If a court finds an officer's use of force unreasonable under the Fourth Amendment, the officer may be held personally liable for damages unless protected by qualified immunity, which shields officers unless they violated "clearly established" law (Harlow v. Fitzgerald, 1982).[2] Federal *criminal* liability is less common, but under 18 U.S.C. § 242, federal prosecutors may pursue charges if an officer willfully deprives a person of their constitutional rights under color of law.[3] Minneapolis police officer Derek Chauvin was convicted under this statute for the murder of George Floyd (USDOJ, 2021); nonetheless, this is exceedingly rare.

State laws also play a role in regulating use of force (Stoughton et al., 2021). The vast majority of states do so via statute. These statutes provide legal defenses and justifications for police officers who use force, varying somewhat by jurisdiction. The difference between assertive force, used to effectuate arrest or prevent flight, and defensive force, when the subject is noncompliant and poses a physical threat, can be an important distinction in state law. When an officer exceeds their statutory or constitutional authority, officers may be charged under state laws such as assault, battery, or homicide. The relationship between constitutional law and state law can be complicated. Uses of force that violate the Constitution may not violate state law, and those that violate state law may not be explicitly unconstitutional (Stoughton et al., 2021). This being said, the legal framework provided by *Graham*, when less-lethal force is concerned, has been incorporated or referenced in statute in most states. Nearly half of states have incorporated or referenced some part of either *Garner* or *Harris* regarding deadly force (Stoughton et al., 2021). In sum, although there are numerous formal legal standards, they consistently consider constitutional case law, such as *Garner*, as a baseline standard against which police use of force is evaluated. Nonetheless, community members set their own informal,

[2]Nonetheless, indemnification regimes ensure that it is vanishingly rare that they ever would be (Schwartz, 2014).

[3]The willfulness test is the most difficult to fulfill, as most federal courts' interpretations have required an outright disregard for the law, protecting officers deemed to have been acting in good faith.

normative standards.

1.2. Community Expectations Standard

In contrast to the legal standards, the Community Expectations Standard evaluates police use of force based on the perceived legitimacy of officers' actions by the communities they serve. This standard is informal and perception-driven, applying whenever a police use-of-force incident enters public consciousness, regardless of whether it triggers a formal investigation or lawsuit. It addresses not only the legality of force, but also whether the use of force aligns with societal norms and values about justice, dignity, and respect (Stoughton et al., 2021). While there is some overlap with the law, there are important distinctions. Further complicating the community expectations standard, police inaction, through failures to intervene or render aid in a situation, is often scrutinized similarly to a use of force. This is another application of the reasonableness divide through the community expecting of police what they are not necessarily obligated to do (Stoughton et al., 2021).

Stoughton et al. (2021), legal scholars and use-of-force experts, derived the five domains by systematically analyzing the common arguments, justifications, and criticisms that frequently appear in public discourse following high-profile use-of-force incidents. The model is therefore a descriptive framework that aims to capture and organize the recurring themes the public uses to evaluate force, providing a structure to understand the "reasonableness divide." The framework was not developed through traditional psychometric scale development but rather through a qualitative, inductive process. The CES criteria are rooted in personal and vicarious experiences of procedural justice/injustice, as well as broader social narratives of institutional legitimacy.

Stoughton et al. (2021) identify five core dimensions that structure how the public often evaluates police use of force under this standard. Three of these factors are generally critical of police use of force, and two of these are favorable.

1.2.1. Underlying Governmental Interest

Force is assessed based on the underlying government interest that initiated the police-citizen encounter. The legal standard established in *Graham v. Connor* (1989) explicitly identifies the severity of the crime at issue as a core factor in determining the reasonableness of force. The CES similarly assesses force based on the government interest that initiated the police encounter. However, the standards diverge in application, as community members often conflate the justification for the initial police contact with the justification for any subsequent force. For instance, when an encounter is initiated over a minor infraction, as in the case of Eric Garner allegedly selling untaxed cigarettes, many members of the public view significant use of force as disproportionate and unjustified, regardless of the subject's later resistance (Fulton-Babicke, 2018). From this perspective, the government's interest in enforcing a low-level offense is seen as too weak to justify force that risks serious harm. In other words, the underlying offense becomes *dispositive* rather than *informative*.

1.2.2. Avoidability

Both legal and community standards also incorporate a concept of necessity, though it is interpreted very differently. While Stoughton et al. refer to the community standard as "The Necessity of Force," we adopt the term "Avoidability" in our analysis.[4] The legal standard of "objective reasonableness" assesses whether an officer's actions were justified by the circumstances they were confronting, explicitly rejecting the use of hindsight. In contrast, the community often evaluates necessity using hindsight, "perfect information," and a variety of other post-hoc criteria that often equate regrettable outcomes to improper methods. Furthermore, the public increasingly expects officers to de-escalate situations and avoid creating unnecessary risks, a concept known as officer-created jeopardy. While some state laws are beginning to require de-escalation, the constitutional standard does not require officers to use the least intrusive means to legally apprehend someone. In short, the public may ask whether the use of force (or its consequence) was in some way avoidable (even if only in hindsight), and if it was, deem it as unnecessary and unreasonable—in other words, avoidable.

1.2.3. Officer Motivation

One of the clearest points of tension between legal and community standards concerns officer motivation. The Supreme Court established that the legal analysis is one of "objective reasonableness," which explicitly assesses an officer's actions without regard to their underlying intent or motivation. Conversely, officer motivation plays a central role in shaping public judgments under the CES. The public often condemns force, even if legally justified, if it appears to stem from biased, vindictive, or authoritarian motives rather than a sincere commitment to public service. Force is frequently viewed as illegitimate if an officer is perceived to be reacting to a subject's race or non-deferential attitude rather than a legitimate threat, tying into broader public concerns about procedural justice. Similarly, force may also be deemed illegitimate if it originates from a pretextual stop, wherein an officer's official justification (a minor offense) serves as a guise for their actual motive of investigating a more serious crime. By way of example, Stoughton et al. (2021) cite the Sandra Bland case, in which a white officer escalated a simple traffic stop of a Black woman because she challenged his order to put out a cigarette and exit the vehicle; many wondered whether he was motivated by race and/or a demand to submit to his authority, rather than any real interest to uphold the law or maintain officer safety (Klein, 2018).[5]

[4] We do this to draw a sharper distinction between the formal legal standard—which assesses necessity from the perspective of a reasonable officer on the scene and the community expectation. The concept of "Avoidability" better captures the public's common use of hindsight and its focus on whether de-escalation or different tactics could have prevented the use of force altogether.

[5] Wolfe et al. (2024) analyzes the "us versus them" mentality observed within the police subculture. This form of dichotomous thought leads officers to view their work as purely being good versus evil. Officers who engage in this thinking are less likely to support the

1.2.4. Subject's Resistance

Force is sometimes viewed primarily as a consequence of the subject's resistance. The law considers a subject's active resistance to arrest or flight as a key factor justifying force, but it requires that force to be proportional. In contrast, community members sympathetic to the police may argue that *any* force is justified by an offender's resistance. This perspective treats the issue as a binary distinction where any non-compliance is seen to justify the officer's actions, often overlooking the legal requirement of proportionality (Stoughton et al., 2021). People often confuse two separate questions: "Was it okay for the police to use force at all?" and "Was the amount of force they used appropriate?" The problem, according to Stoughton et al. (2021), is that when *some* resistance justifies *some* force, many community members justify *any* level of force. If this community view is widespread, then we may expect no covariation between reasonableness appraisals and subject resistance.

1.2.5. "Highly Dangerous" Environments

While the law allows consideration of the "totality of the circumstances" to evaluate risk, some community members perceive many circumstances as inherently dangerous, even without objective evidence. This perspective justifies force not because it was ideal, but because the job is hazardous and mistakes are excusable. Factors such as the presence of bystanders or the time of day might be perceived as making a situation highly dangerous, leading to deference to officers in situations where a strict legal analysis of an immediate threat might not support the use of force. Stoughton et al. (2021) specifically identify unarmed, non-combative crowds as an example of an ambiguous source of "danger" that some community members have used to justify police use of force (Clark et al., 2017). The quotes used in the phrase *"highly dangerous" environment* emphasize its more rhetorical nature over objective evidence of danger.

1.3. Prior Evaluations of the CES

While relatively little research has directly tested the CES model, the early results tend to support it. Fridell and Marier (2023) experimentally tested two CES factors by manipulating a written vignette shown to a national sample of 3,601 U.S. adults. The study examined the *underlying governmental interest* by varying the seriousness of the offense that precipitated the police encounter (a broken taillight vs. a felony warrant) and *officer motivation* by varying the suspect's race (Black vs. White). The results provided strong support for the CES model. Participants were significantly more likely to find the officer's use of deadly force reasonable when the precipitating event was serious versus not

use of de-escalation tactics, procedural justice, and the prioritization of self-control. These officers are more likely to perceive threats after being presented with body camera footage. This research ties into broader themes of Warrior vs. Guardian orientations (McLean et al., 2020), which somewhat reflects notions of due process vs. crime control (DeLisi, 2011). In short, officers' motivations may guide both their conduct and the public's interpretation of it.

serious, irrespective of equivalent suspect behavior that followed. The study also found that force was deemed significantly less reasonable when used against a Black suspect compared to a White suspect, presumably due to inferences about the officer's *motivations* and racial bias.

Richardson and Fridell (2025) assessed a broad range of the CES perspectives using a national sample of 1,200 U.S. adults who evaluated use-of-force vignettes. The study experimentally manipulated two aspects of the *avoidability* dimension: perfect information/hindsight (providing after-the-fact knowledge of the suspect's dangerousness) and officer-created jeopardy. It also measured other CES criteria using survey questions about participants' perceptions of the incident (although these were not experimentally manipulated). The study found a significant "reasonableness divide," with participants' assessments aligning with the legal standard only 54.6% of the time—marginally better than chance. The CES perspectives were powerful predictors of reasonableness assessments, explaining nearly 65% of the variance in participant judgments. While legal reasonableness impacted assessments, it had the weakest impact of all significant predictors. In other words, the legally *irrelevant* factors were more strongly related than the legally relevant factors.

2. Current Study

While a substantial body of research has documented the demographic and attitudinal predictors of public opinion on police use of force, these studies have not offered a cohesive theoretical framework to explain these evaluations. The Community Expectations Standard attempts to fill this void. While two studies have supported some of its hypotheses, only a subset of the standards were tested using experimental methods. Furthermore, no study to date has considered the use of non-lethal force, which is less serious and controversial but nonetheless *far* more common. The present study is therefore the first to systematically and empirically evaluate all five domains of the Community Expectations Standard in a single, unified experiment and to consider its applicability to non-lethal force. To accomplish this, we employ a factorial survey experiment that manipulates each of the community standards while maintaining the standard of legal reasonableness across conditions. This methodology allows for the isolation of the independent causal effect of each CES domain on public perceptions of reasonableness—an advantage not afforded by observational or correlational research.

The CES model, if validated, may provide a systematic framework to help explain the tension between the law and public opinion. By identifying and explaining the specific criteria the public uses to evaluate force, the model can help policymakers and police officials understand public evaluations, facilitate more effective communication, and ultimately work toward developing policies and training that can bridge the gap between legally justified conduct and police legitimacy.

3. Methods

A national sample of adults completed a survey in which they evaluated the reasonableness of a police officer's use of force as described in a hypothetical vignette in the style of a local news story taking place at a convenience store. Respondents were randomly assigned to one of two conditions across five variables, including underlying governmental interest, avoidability, officer motivation, subject resistance, and a "highly dangerous" environment. Participants' demographic information was also collected. Details regarding the sample, experimental procedures, measures, and analysis are provided below.

3.1. Sample

Participants of this study were a national sample of 1,955 American adults selected from the population of individuals who have signed up to participate regularly in online surveys administered by Cint Theorem (formerly Lucid Theorem). Cint Theorem provides nationally reflective samples of U.S. adults to academic and market researchers. Participants agree to regularly participate in surveys in exchange for incentives (e.g., gift cards). Cint Theorem's recruitment method is similar to those of other well-recognized survey firms, including Qualtrics, Dyanata, CCES, and SSI (Norris et al., 2023). Such samples are more demographically and politically representative than those obtained via Facebook or Mechanical Turk (Boas et al., 2020). Quota-based convenience samples consistently produce treatment effects and effects sizes that are equivalent to more costly nationally-representative population-based samples (Mullinix et al., 2015). Power analysis based upon pilot results indicated that a sample size of approximately 1,800 respondents was necessary to detect a small (.15) main effect.

Table 1 reports the sample demographics. The sample was 79.7% White (of any ethnicity, including Hispanic/Latino), 14.1% Black, and 13.2% Hispanic (of any race), while 48.1% of respondents were male. This compares to national 2020 census benchmarks, where 61.6% of Americans were White (of any ethnicity), 12.4% were Black, 18.7% were Hispanic (of any race), and 49.1% were male. We speculate the somewhat lower sampling of Hispanics/Latinos may be due to the politically charged topic (policing) and the inauguration of Donald J. Trump and his implementation of "mass deportation" policies, given that they were contemporaneous with the survey's distribution. A total of 2,614 survey invitation responses were received. Responses were eliminated from the study sample for any of several reasons: declining to participate after informed consent (119); non-completion after opt-in (222); failing either of two attention checks (255); and/or if reCAPTCHA estimated the "respondent" was more likely to be a bot than a human (67).[6] This reduced the final sample to 1,955 respondents.

[6]These are not mutually exclusive categories.

Table 1: Descriptive Statistics

Variable	Mean (SD) or %
Reasonableness Appraisals	
Not at all Reasonable	8.08%
Very Unreasonable	9.82%
Somewhat Unreasonable	15.45%
Somewhat Reasonable	27.42%
Very Reasonable	22.51%
Completely Reasonable	16.73%
Experimental Manipulations	
Government Interest	51%
Avoidability	49%
Motivation	50%
Resistance	50%
"Dangerous" Env.	49%
Personal Characteristics	
Male	48.1%
Hispanic/Latino	13.2%
Age	31.86 (17.85)
Race	
Asian	3.96%
Black	14.11%
Native American	1.95%
Pacific Islander	0.33%
White	79.65%
Education	
Some H.S. or Less	2.25%
H.S. Graduate	21.98%
Post-H.S. Vocational	4.35%
Some College	18.24%
Associate's	15.11%
Bachelor's	23.98%
Master's or Professional	11.32%
Doctorate	2.77%
Income	
$14,999 or less	10.95%
$15,000 - $34,999	22.10%
$35,000 - $64,999	28.44%
$65,000 - $124,999	18.93%
$125,000 - $249,999	17.60%
$250,000 or more	1.99%
Partisanship	
Strong Democrat	10.91%
Lean Democrat	28.07%
Independent	19.85%
Lean Republican	29.88%
Strong Republican	11.29%

Note. N varies by variable, from 1,824 to 1,955, due to missing data.

3.2. Procedures

Respondents were randomly assigned to versions of a vignette that varied (1) the underlying governmental interest, (2) avoidability, (3) officer motivation, (4) suspect resistance, and (5) "highly dangerous" environments—a 2×2×2×2×2 factorial design. The full list of conditions, and their corresponding domain within the Community Expectations Standard, is illustrated in Table 2. The full vignette with manipulations appears in Appendix A. The vignette was pre-screened by three use-of-force experts and a legal scholar to ensure it was consistent with both the law (legally reasonable across conditions) and the theory (faithful to the five domains articulated by Stoughton et al. 2021), including one of the theory's authors. In all scenarios, the police officer responded by using a Taser, causing serious but non-lethal injury. We hypothesized that an officer's use of force would be considered more reasonable in Condition 2 for all factors.

Table 2: Experimental Conditions/Manipulations

CES Domain	**Condition 1**	**Condition 2**
Underlying governmental interest	Suspect attempted to to use what the clerk thought was a counterfeit $10 bill.	Suspect broke open an ATM.
Avoidability	Officer kicked down a door in an attempt to apprehend a suspect locked in a bathroom.	Officer conversed with the suspect in order to persuade him to voluntarily step out of a locked bathroom.
Motivation	Officer is characterized as not tolerating hoodlums, thugs, or disrespect.	Officer is characterized as deeply committed to fair and effective policing.
Suspect resistance	Suspect actively struggled and attempted to pull away during an arrest.	Suspect lunged toward the officer and attempted to grab his Taser.
"Highly Dangerous" Environment	The suspect was the only person in the store.	The suspect's friends arrived and started filming the incident.

3.3. Measures

The dependent variable is a measure of perceived *reasonableness* of police use of force. After reading the vignette, respondents were asked to rate how reasonable the use of force was on a six-point scale, from "Not at all Reasonable" (=1) to "Completely Reasonable" (=6).

The core independent dummy variables reflect the experimental manipulations of the five CES factors as described above. These are objective, dichotomous experimental manipulations, rather than respondents' subjective evaluations of the event. As illustrated in Table 1, the experimental manipulations were well balanced across conditions. Additionally, the survey captured "embedded data," which refers to certain respondent characteristics that were previously collected by the survey firm but are not included as survey questions. Therefore, we also measure and model respondent gender (male=1), race (dichotomized into a dummy variable where Black=1), ethnicity (a dummy variable, Hispanic=1), age (continuous in years), education (on a scale where 1=Less than a high school degree and 8=A Doctorate degree), household income (a 6-point ordinal scale, where the lowest value indicated an income of less than $15,000 per year, and the highest indicated an income over $249,999 per year), and partisanship (a 5-point ordinal scale where 1=Strong Democrat and 5=Strong Republican).

3.4. Analytic Strategy

Analysis begins with comparisons of means tests. Independent samples t-tests are reported to evaluate the causal effect of each factor on reasonableness appraisals. With an experimental design and random assignment, straightforward comparison-of-means tests provide strong evidence of independent causal effects.[7]

Next, we ran Bayesian regression analyses that model the variables simultaneously. This approach was chosen over traditional frequentist methods due to several key advantages that align with our research goals (Mourtgos, 2025; Kruschke and Liddell, 2018). First, Bayesian analysis provides a more intuitive interpretation of uncertainty. Instead of p-values and confidence intervals, the analysis yields a full posterior distribution for each parameter. From this distribution, we can calculate 95% credible intervals, which represent a range of plausible values for an effect with a specific probability. This allows for direct probabilistic statements about the magnitude and direction of each factor's influence. Second, this framework allows for the direct quantification of evidence for or against a hypothesis. For instance, rather than relying on a binary significant/non-significant decision, we can calculate the posterior probability that a coefficient is positive. This provides a more nuanced understanding of the certainty of each effect, moving beyond arbitrary p-value thresholds.

Regression analyses were conducted in R using the *brms* package. To ensure stable and regularized estimates, we used weakly informative priors for the regression coefficients. Model convergence was assessed by running four independent Markov Chain Monte Carlo (MCMC) chains, each for 4,000 iterations with a warmup period of 1,000 iterations. We confirmed convergence by ensuring the potential scale reduction factor ($\hat{R}$) was approximately 1.0 for all parameters and by visually inspecting the trace plots.

[7]Non-parametric Mann-Whitney U tests converged with the t-tests reported here.

We built two regression models. The first models reasonableness on all five CES factors to simultaneously estimate the independent effects of each experimental manipulation, controlling for all others. The second model adds several personal characteristics to explore how they independently predict reasonableness assessments after accounting for CES factors inhered to the police–citizen encounter.

4. Results

Table 3 reports the results of independent samples t-tests. Each of the CES experimental manipulations predicts significantly higher perceptions of police use-of-force reasonableness at $p < .05$, except *officer motivation*, which a non-significant trend ($.05 < p < .10$). Cohen's d values indicate that *government interest* ($d = .35$) and *subject resistance* ($d = .42$) demonstrate substantive effects, whereas the other factors' effects are small to negligible ($.09 < d < .13$). In short, the experimental results show that the public perceives police force as more reasonable when: (1) the underlying governmental interest is greater (that is, when the underlying crime is more severe, even when it remains non-violent); (2) the officer avoids escalation (that is, the officer uses time and dialogue rather than rushing to effectuate an arrest); (3) the subject demonstrates somewhat higher resistance (here, attempting to disarm the officer's taser as opposed to actively resisting arrest); and (4) the environment is "more dangerous" (in this case, crowds are gathering, creating an ambiguous "threat"). On the other hand, evidence that the officer's intentions matter—that he is motivated by fair and effective policing rather than more authoritarian motives—demonstrates a smaller effect, if any.

Table 3: Results of Two-Sample t-tests and Effect Sizes

	Cond. 1		Cond. 2			
Variable	Mean	SD	Mean	SD	t-statistic	Cohen's d
Governmental Interest	3.71	1.52	4.22	1.38	7.84***	.35
Avoidability	3.87	1.46	4.07	1.48	2.93**	.13
Officer Motivation	3.90	1.48	4.03	1.46	1.93†	.09
Subject Resistance	3.66	1.47	4.27	1.41	9.35***	.42
"Dangerous" Environment	3.89	1.48	4.05	1.46	2.39*	.11

Note: Significance levels are denoted as follows: $^{\dagger}p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Next, we modeled all five CES factors simultaneously in a Bayesian linear regression model, reported in Table 4. This permits the estimation of each factor's *independent* effect, controlling for all others; non-experimental control variables are omitted at this stage to continue exploring causal effects from random assignment, without consideration for non-experimental variables.

The largest effects are observed for *governmental interest* (b = .513) and *subject resistance* (b = .605), where the posterior probability that these effects are positive is greater than 99.9%. The effect of avoidability was smaller but notable (b = .126, CI [0.002, 0.250]), with a 97.7% posterior probability of a positive effect. The effect of *officer motivation* and *"highly dangerous" environment* were more uncertain; while the mean estimates were positive, the 95% credible interval included zero (*motivation*: b = 0.097, 95% CI [-0.031, 0.224]; *"highly dangerous" environment*: b = .097, 95% CI [-0.004, 0.245]), and the probabilities of positive effects were 93.3% and 97.1%, respectively.

Table 4: Bayesian Linear Regression Model Predicting Reasonableness Ratings

Predictor	*b*	*SD*	95% Credible Interval	*P(Direction)*
Governmental Interest	0.513	0.065	[0.385, 0.640]	>.999
Avoidability	0.126	0.064	[0.002, 0.250]	.977
Officer Motivation	0.097	0.064	[-0.031, 0.224]	.933
Subject Resistance	0.605	0.065	[0.478, 0.733]	>.999
"Dangerous" Environment	0.121	0.063	[-0.004, 0.245]	.971
(Intercept)	3.236	0.075	[3.090, 3.380]	
Sigma (σ)	1.415	0.022	[1.373, 1.460]	

Note. N = 1955. b represents the posterior mean and *SD* represents the posterior standard deviation of the coefficient. The 95% Credible Interval is the equal-tailed interval of the posterior distribution. *P(Direction)* is the posterior probability that the effect is in the same direction as the posterior mean. Model convergence was confirmed with $\hat{R}$ values of 1.00 for all parameters.

These results are visualized in Figure 1, which illustrates the posterior distributions of the regression coefficients as well as the value of reasonableness predicted by each treatment condition, after controlling for all other factors. The predicted values in panel B suggest that not only do high governmental interest and high subject resistance predict more reasonable appraisals, but that *low* governmental interest and *less* subject resistance produce more *un*reasonable appraisals. In summary, the figure clearly illustrates that *governmental interest* and *subject resistance* are the primary drivers of reasonableness ratings in the experiment, while *necessity* and *"highly dangerous" environment* have a smaller positive impact and *officer motivation* had little to no effect.

Finally, we estimated a model that adds several sociodemographic covariates to the linear regression model, including gender, race, ethnicity, age, education, income, and partisanship (where higher values reflect more conservative/Republican alignment). The results appear in Table 5. Several observations are noteworthy. As before, *governmental interest* and *subject resistance* demonstrate the strongest effects among the experimental manipulations. *"Highly dangerous" environment* demonstrated the largest drop in effect, both in terms of its regression coefficient and the posterior probability. This may be due to the fact that one or more of the sociodemographic factors moderate this manipulation. In contrast to much prior research, race (b = −.120, CI [-0.319,

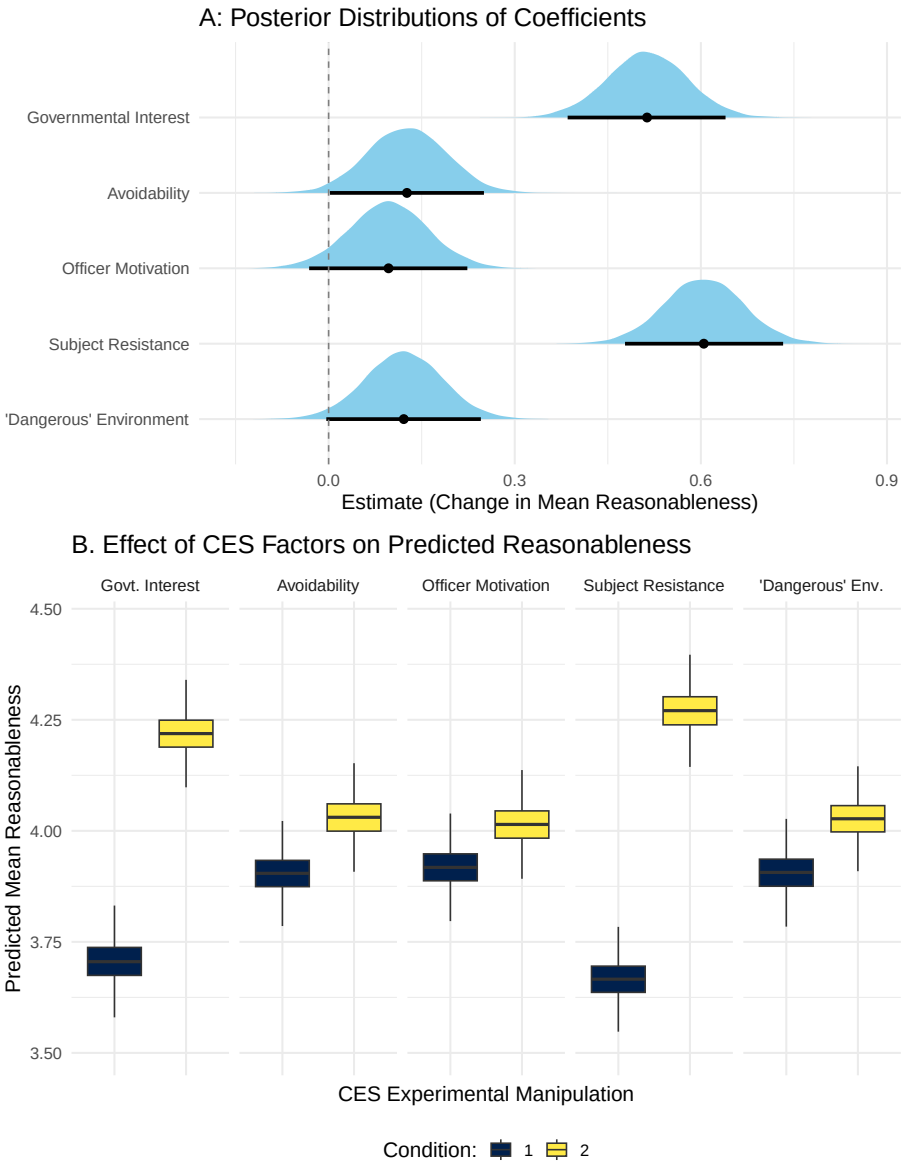

Figure 1: Posterior Distributions of Coefficients and Predicted Values

0.071]) and ethnicity ($b = -.042$, CI [-0.256, 0.168]) are rather poor predictors of perceptions of use-of-force reasonableness, where 0 falls well within their 95% credible intervals. (Male) gender ($b = .126$) increases reasonableness appraisals, although the 95% credible interval just barely contains zero (CI [-0.008, 0.261]). Age ($b = .008$, CI [0.005, 0.012]), income ($b = .057$, CI [0.004, 0.110]) and (conservative) partisanship ($b = .256$, CI [0.203, 0.310]) independently predict higher reasonableness appraisals. Overall, the variables with the strongest substantive effects on perceptions of reasonableness are *underlying governmental interest*, *subject resistance*, and *partisanship*.

Table 5: Full Bayesian Linear Regression Model with Covariates Predicting Reasonableness

Predictor	*b*	*SD*	95% Credible Interval	*P(Direction)*
Experimental Factors				
Governmental Interest	0.608	0.090	[0.433, 0.786]	>.999
Avoidability	0.126	0.064	[-0.003, 0.252]	.973
Officer Motivation	0.097	0.064	[-0.029, 0.223]	.935
Subject Resistance	0.729	0.091	[0.547, 0.908]	>.999
"Dangerous" Environment	0.093	0.064	[-0.033, 0.219]	.924
Covariates				
Male	0.126	0.069	[-0.008, 0.261]	.966
Black	-0.120	0.099	[-0.319, 0.071]	.889
Hispanic	-0.042	0.108	[-0.256, 0.168]	.654
Age	0.008	0.002	[0.005, 0.012]	>.999
Education	0.021	0.020	[-0.017, 0.059]	.850
Income	0.057	0.027	[0.004, 0.110]	.983
Partisanship	0.256	0.028	[0.203, 0.310]	>.999
Model Parameters				
(Intercept)	1.851	0.166	[1.525, 2.181]	—
Sigma (σ)	1.369	0.023	[1.325, 1.414]	—

Note. N = 1,812. *b* represents the posterior mean; *SD* is the posterior standard deviation. The 95% Credible Interval is the equal-tailed interval of the posterior distribution. *P(Direction)* is the posterior probability that the effect is in the same direction as the posterior mean. Model convergence was confirmed with $\hat{R}$ values of 1.00 for all parameters.

5. Discussion

A substantial body of research has explored how personal identity influences citizens' appraisal of police use of force. This research has consistently found that demographic characteristics and partisanship frequently correspond to these appraisals. Yet, there has been surprisingly little theorizing about these associations. Stoughton et al. (2021) have proposed a model that attempts to

identify criteria that the public applies to evaluations of police use of force in their "Community Expectations Standard." This CES is an additional, informal standard to which police are held, supplementing legal and administrative standards. This study sought to test the CES, building on two prior studies by simultaneously examining all five proposed domains of evaluative criteria in an experimental study using a national sample of U.S. adults. Importantly, it did so using factors that varied on CES criteria but were nonetheless legally reasonable, in order to explore a "reasonableness divide" in "lawful but awful" use of force.

Our study found strong evidence for independent causal effects for *underlying governmental interest*, *avoidability*, *subject resistance*, and a *"highly dangerous" environment*. That is, the public views police use of force as more reasonable when it involves a more serious crime (even when non-violent), when the officer exercises patience and de-escalation rather than rushing to effectuate an arrest, when the suspect is demonstrating higher levels of resistance, and when the officer faces more ambiguously "dangerous" situations in the presence of crowds. There was a non-significant trend toward a positive association between the officer's *motivations* and perceptions of his use of force. Although all of the vignette conditions consisted of conditions that were *prima facie* reasonable, the experimental manipulations demonstrated that community members, on average, consider police use of force to be less reasonable based on these specific case characteristics. Thus, the police are held not only to a legal standard, but also to Community Expectations Standard.

The largest effects are observed for *underlying governmental interest* and *subject resistance*. Importantly, these two evaluative criteria are also those with analogs in the law. The severity of the initial crime had a large and significant effect on public perception of reasonableness. This mirrors the *Graham* factor permitting the consideration of "the severity of the crime at issue." The level of the suspect's resistance was the strongest predictor of perceived reasonableness. This aligns with *Graham*'s consideration of "whether the suspect was actively resisting arrest or attempting to evade arrest by flight." This suggests that while there is a "reasonableness divide," the public's judgment is not entirely divorced from legal principles. Community members are most influenced by the same core situational factors that the law deems most important. Nonetheless, there are distinctions. Under *Graham*, the seriousness of the crime is generally salient to the consideration of offenses involving weapons, threats, or violence, as proxies for dangerousness. Here, both experimental conditions involved non-violent property crimes, suggesting that community members are attuned to the seriousness of the precipitating event not because it is necessarily an indication of dangerousness, but because they are sensitive to the balance between the state's legitimate crime control interests, and the public's civil liberties.

While the legal standard in *Graham* treats active resistance as sufficient justification for force, our results show that community members are more discerning, deeming the use of a Taser significantly more reasonable in response to aggressive resistance compared to less-serious active resistance. This finding presents a direct challenge to a key hypothesis within the CES framework.

Stoughton et al. suggest that many community members adopt a binary perspective on non-compliance, where any resistance justifies any subsequent force. If such a viewpoint were widespread, the specific *level* of resistance in our experiment should have had a negligible effect on reasonableness ratings. The strong, substantive effect we observed suggests that this all-or-nothing view of resistance is either not widespread, or its application is conditioned by other situational or respondent characteristics.

Officer motivation had little to no effect across models. Perceptions of an officer's motives are likely inferred and heavily influenced by a respondent's pre-existing trust in police, which is strongly tied to race and partisanship. These findings contrast with prior research on CES, which operationalized *officer motivation* via the manipulation of suspect race, suggesting that community members infer racial animus in interracial encounters (Fridell and Marier, 2023). The present study explicitly excluded the race of the officer and suspect, opting instead to use a statement through the Chief of Police alluding to the officer's motives (essentially, characterized with a professional due process vs. authoritative crime control orientation). Our subtler manipulation, the weak results, and the contrast to prior research suggest that the role of *officer motivation* on public appraisals of police use of force deserves more empirical attention.

The large, consistent effect of *underlying governmental interest* is consistent with prior experimental research evaluating the influence of the precipitating incident on subsequent appraisals of force (Fridell and Marier, 2023), suggesting that the state's relative interests are a core, foundational criteria in the public's evaluation of coercive force serving those interests. Nonetheless, in prior non-experimental research, perceived offense seriousness had no bearing on use-of-force appraisals (Richardson and Fridell, 2025). And while we found large effects for *subject resistance*, this again contrasts with prior non-experimental null findings for perceived suspect resistance (Richardson and Fridell, 2025). In fact, the best-performing factors in the non-experimental research—*officer motivation* and *proportionality* (an aspect of necessity/avoidability)—were among the *weakest* factors in this experiment. This illustrates the way that some CES factors may be relatively objective characteristics of an event, up for little dispute, while other CES factors are attenuated or moderated through the participant's personal and political identity, as we discuss further in the next section.

5.1. Theoretical Implications

The weaker relationships found in our study contrast with the rather strong relationships found in prior correlational research that involved asking about citizens' perceptions of an event, rather than experimental manipulation of it (Richardson and Fridell, 2025). When considered in conjunction with the original articulation of the Community Expectations Standard (Stoughton et al., 2021), we ask whether these standards are best understood as *characteristics of the event*, or *characteristics of the respondent*. For example, are "highly dangerous" environments best understood as ambiguous threats which officers actually face, or post-hoc rationalizations used by specific segments of society?

If the latter, which segments? Is it possible that *some* CES criteria are event characteristics, while *others* are respondent characteristics—and how do these interact?

Our results suggest a refinement to the CES model by proposing a more explicit distinction between event-based and observer-based criteria. The factors with the strongest and most consistent effects in our experiment—*underlying governmental interest* and *subject resistance*—function as event-based characteristics. Like their analogs in constitutional law, these are concrete, relatively objective features of the encounter that produce more consistent judgments across different groups. Their strength suggests that public evaluations of force are not entirely subjective but are grounded in shared, foundational principles of proportionality that the public weighs heavily when appraising an officer's actions.

Conversely, other CES domains appear to function as observer-based criteria. These factors, such as the perception of a *"highly dangerous" environment* and *officer motivation*, are not objective features of the event but are subjective interpretations mediated by the preexisting beliefs, identities, and political leanings of the person evaluating the encounter. For instance, the effect of the *"highly dangerous" environment* manipulation lost predictive power once respondent characteristics were introduced, suggesting the perception of "danger" in this context is less about the event and more about the observer's worldview. This is supported by literature on the ideological asymmetry of threat sensitivity, which suggests that political conservatism is robustly linked to a "generalized susceptibility to experiencing threat or anxiety in the face of uncertainty" (Wilson 1973, p. 259): conservatives are, on average, more psychologically and physiologically sensitive to threatening stimuli than liberals (Jost et al., 2017). While Stoughton et al. (2021) suggest that personal orientations are salient to these community standards, we propose this distinction be made more explicit and central to the model. The CES may be better conceptualized not as a flat model of five domains, but as a framework containing two distinct types of criteria that operate differently in shaping public opinion.

In the current political climate, partisanship may function as a master heuristic for interpreting police actions, potentially overriding other historically paramount factors (Boudreau et al., 2020). In contrast to a great deal of prior research, race and ethnicity had no bearing on reasonableness appraisals in our models; instead, partisanship was one of the most powerful predictors. Partisanship does not just predict attitudes; it appears to shape the perception of the facts of the event itself. This aligns with broader theories of motivated reasoning, where individuals process information in ways that confirm their partisan identities (Bolsen et al., 2014). For the study of policing, this implies that legitimacy is increasingly tied not just to the specific actions of the police, but to the broader political narratives in which those actions are situated and interpreted (Anderson and Just, 2013).

That partisanship functions as a master heuristic aligns with prior work that has found that conservative ideology is linked to consistently high approval of force, while liberal ideology leads to more variable judgments that distinguish

between reasonable and excessive scenarios (Navarro and Hansen, 2023). Other studies have pinpointed that ideological constructs like Right-Wing Authoritarianism (RWA) and Social Dominance Orientation (SDO) are unique predictors of support for excessive force, while legitimacy predicts support for reasonable force (Gerber and Jackson, 2017). The divergence or convergence of opinion can even depend on the type of error an officer makes; for instance, liberals are uniquely punitive of 'false-positive' errors (shooting an unarmed suspect) but converge with conservatives in their tolerance of 'false-negative' errors (Patil, 2018). In a highly polarized landscape, partisanship itself, as a core social identity, may now be operating as a more dominant and accessible cognitive shortcut than these more nuanced ideological dimensions, powerfully shaping an observer's interpretation of the facts from the outset. Future research must continue to disentangle whether the CES criteria are best understood as objective features of encounters or as subjective frames that are themselves a product of citizen identity and political polarization.

5.2. Policy Implications

The findings of this study offer several important policy implications for law enforcement agencies, policymakers, and community leaders seeking to bridge the "reasonableness divide" and enhance police legitimacy. First, police training might be expanded beyond the legal standard of objective reasonableness to explicitly incorporate the core tenets of the Community Expectations Standard. Our results show that the public is highly sensitive to the *underlying governmental interest* and the *avoidability* of force. Therefore, training curricula should emphasize de-escalation tactics and strategic disengagement, particularly in encounters stemming from minor, non-violent offenses. POST boards could mandate scenario-based training that teaches officers de-escalation and assessment techniques, such as the Police Executive Research Forum's demonstrably effective ICAT (Integrating Communications, Assessment, and Tactics) training (Engel et al., 2020). Officers should be trained to understand that even if force is legally permissible, its use in a low-level encounter will likely be viewed as illegitimate by the community, eroding public trust.

Second, law enforcement agencies may consider revising their use-of-force policies to be more restrictive than the legal minimum standards. This administrative policy reform would involve a comprehensive internal review of existing guidelines, potentially guided by model policies from professional organizations like the Police Executive Research Forum or International Association of Chiefs of Police. Policies that codify a duty to de-escalate, limit the types of force permissible for non-violent resistance, and place a higher premium on the sanctity of life can help align agency standards with community expectations. By proactively adopting stricter internal standards, departments can demonstrate a commitment to procedural justice that goes beyond mere compliance with the law.

Third, communication strategies following a critical use-of-force incident must address the factors that matter most to the public. Beyond stating that an action was "lawful," official statements should clearly articulate the severity

of the initial offense and the specific nature of the subject's resistance, as these are the most powerful drivers of public perception. Transparency regarding these factors can help contextualize an officer's actions in terms the public understands and weighs heavily. Any misalignment between police behaviors and public expectations might be reconciled not only by changing police behavior but also by shifting public expectations. To that end, community members must be educated about the statutory and constitutional legal standards for use of force, as well as their rationale. These standards must be consistently identified in after-action communications, and agencies must be transparent about any inconsistencies or vagaries in the law. Although all fifty states establish civics standards, few if any curricula address the government's bounded authority to use physical coercion. This is a critical gap, given that police are the most public-facing government body most citizens encounter. Out of necessity, if not desire, law enforcement agencies must take on this educational role, even if they are imperfect messengers facing skeptical audiences.

Finally, the powerful influence of partisanship highlights the limits of case-specific explanations and the critical need for long-term trust-building initiatives. Since political identity so strongly colors perceptions of police actions, agencies must invest in developing positive, non-enforcement relationships across all communities. A deep reservoir of public trust, built on a foundation of consistent procedural justice and community engagement, is the most effective bulwark against the polarizing narratives that often dominate the discussion around police use of force.

5.3. Limitations and Future Directions

Several limitations should be considered when interpreting the findings of this study. First, our research relied on a hypothetical, text-based vignette to assess public perceptions. While this experimental method provides a high degree of internal validity by allowing us to isolate the causal effects of specific factors, it lacks the external validity of real-world incidents. Written scenarios cannot fully capture the dynamic, emotionally charged, and often ambiguous nature of actual use-of-force encounters. It is plausible that respondents would evaluate a real event, particularly one viewed on video, differently than the sterile description provided in the survey.

Second, the operationalization of some Community Expectations Standard domains may have contributed to their relatively weak effects. The manipulation for a *"highly dangerous" environment*, which involved friends arriving to film the incident, was designed to be an ambiguous threat. Its ambiguous manipulation, coupled with its ambiguous influence does not resolve whether the public *on average*, views innocuous "threats" such as crowds as "dangerous." Similarly, *officer motivation* was conveyed indirectly through a statement from a police chief about the officer's character. This contrasts with prior research that manipulated suspect race to proxy for officer motivation and found a strong effect, suggesting that our more direct but less visceral manipulation may not have been as salient to respondents. The weak empirical findings could also be a result of the limited analytic power.

Third, the tension between event characteristics and respondent characteristics described in the theoretical implications section above could have been resolved had we incorporated manipulation checks into the experiment. To that end—as well as the one above—we have a replication and extension underway to resolve and clarify this theoretically important distinction.

Finally, while online opt-in samples like the one used here are common in social science research and have been shown to produce reliable treatment effects, they are not true probability samples. As such, while the causal relationships identified in the experiment are robust, caution should be exercised in generalizing the precise distribution of opinions to the entire U.S. population. Future research should build on these findings by testing the CES model using more dynamic stimuli, such as video vignettes, and exploring its applicability across a wider range of use-of-force scenarios to enhance the generalizability of the results.

6. Conclusion

A persistent "reasonableness divide" separates the legal standards governing police use of force from the court of public opinion. This study sought to bring empirical clarity to this divide by systematically testing the five domains of the Community Expectations Standard, a model designed to articulate the criteria citizens use when evaluating police conduct. Through a national survey experiment, we found that the public's judgment of force is complex, guided by a mix of situational factors and the personal characteristics of the observer.

Our findings reveal that the most powerful influences on public perception are the CES domains with direct analogs in constitutional law: the *underlying governmental interest* and the degree of *subject resistance*. This suggests that the public's evaluative framework is not entirely divorced from legal principles; citizens, like the courts, weigh the severity of the crime and the suspect's behavior heavily when appraising an officer's actions. However, the public's application of these standards is more nuanced, distinguishing between levels of non-violent crime and forms of resistance in ways the legal standard may not.

At the same time, this research highlights the substantial influence of partisanship as a lens through which police actions are viewed. Some CES factors, such as the perception of a "highly dangerous" environment, appear less as objective characteristics of an event and more as subjective interpretations filtered through a political worldview. Indeed, partisanship proved to be a more potent predictor of reasonableness judgments than race or ethnicity, suggesting that in our current polarized climate, political identity may be the primary heuristic for making sense of police use of force.

Ultimately, the Community Expectations Standard provides a valuable framework for understanding why legally justifiable force is often seen as illegitimate by the public. Bridging the "reasonableness divide" will require more than just ensuring police actions are lawful; it will demand that police practices also align with the public's normative standards of fairness, necessity, and proportionality. It will also require a public that understands the legal justifications for,

and limits of, the state's right to use force. By identifying the specific criteria that shape public perception, this research can help inform police training, policy, and communication strategies aimed at rebuilding trust and enhancing the legitimacy of law enforcement in a democratic society.

Appendix A. Experimental Vignette

Man Hospitalized After Police Response

by Joseph Anderson, Staff Writer, Daily News

Earlier today, police were called to a convenience store for a non-violent crime, but by the end of it, one man would be end up in an area hospital after police were called to the scene.

Police responded after a 911 call from the store clerk, who said the man, Jacob Jones, [GOV'T INTEREST: attempted to use what the clerk thought was a counterfeit $10 bill/broke open an ATM].

After a brief verbal altercation with the responding officer, Jones locked himself inside the store's bathroom. By this time, ["HIGHLY DANGEROUS" ENVIRONMENT: Jones was the only customer in the store/several of Jones' friends had arrived and started filming the incident]. The officer [AVOIDABILITY: kicked down the door and rushed in to apprehend Jones/conversed with Jones and eventually persuaded him to step out].

At that point, police say, Jones [RESISTANCE: lunged toward the officer and attempted to grab his Taser/actively struggled and attempted to pull away during an arrest attempt]. The officer deployed his Taser, and Jones fell, struck his head on the ground, and lost consciousness. The clerk reported seeing blood coming out of the man's ears, which experts we spoke with said could be consistent with a traumatic brain injury sustained in the fall.

The man was rushed to Memorial Hospital by paramedics, where he remains in critical condition. Police have not yet released any details about the officer involved, but the incident was captured on the store's security cameras. That footage has not yet been made public.

The Chief released a statement, saying, "I've known this officer personally for many years. I know his character well. [OFFICER MOTIVATION: He does not tolerate hoodlums, thugs, or disrespect./He is deeply committed to fair and effective policing.]"

The officer has been placed on paid administrative leave while the investigation continues.

References

Anderson, C.J., Just, A., 2013. Legitimacy from above: The partisan foundations of support for the political system in democracies. European Political Science Review 5, 335–362. doi:10.1017/S1755773912000082.

Baker, M.A., Bacharach, V.R., 2017. Police officer-civilian confrontations caught on camera: The influence of contextual frames on judgements of excessive force. American Journal of Criminal Justice 42, 683–697. doi:10.1007/s12103-017-9387-5.

Bittner, E., 1990. Aspects of police work. Northeastern University Press.

Boas, T.C., Christenson, D.P., Glick, D.M., 2020. Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics. Political Science Research and Methods 8, 232–250.

Bolsen, T., Druckman, J.N., Cook, F.L., 2014. The influence of partisan motivated reasoning on public opinion. Political Behavior 36, 235–262.

Boudreau, C., MacKenzie, S.A., Simmons, D.J., 2020. Partisanship and Public Perceptions of the Police. Working Paper. Department of Political Science, University of California, Davis.

Bradford, B., Milani, J., Jackson, J., 2017. Identity, legitimacy and "making sense" of police use of force. Policing: an international journal 40, 614–627.

Braga, A.A., Winship, C., Tyler, T.R., Fagan, J., Meares, T.L., 2014. The salience of social contextual factors in appraisals of police interactions with citizens: A randomized factorial experiment. Journal of Quantitative Criminology 30, 599–627. doi:10.1007/s10940-014-9216-7.

Carter, J.S., Corra, M., 2016. Racial resentment and attitudes toward the use of force by police: An over-time trend analysis. Sociological Inquiry 86, 492–511. doi:10.1111/soin.12136.

Carter, J.S., Corra, M., Jenks, D.A., 2016. In the shadows of Ferguson: The role of racial resentment on White attitudes towards the use of force by police in the United States. International Journal of Criminal Justice Sciences 11. URL: `https://ijcjs.com/menu-script/index.php/ijcjs/article/view/178`.

Choi, J., 2019. Comparing global and situational support for police use of force across immigrant generations and native-born Americans. Policing: An International Journal 42, 1038–1051. doi:10.1108/PIJPSM-02-2019-0025.

Clark, M.D., Bland, D., Livingston, J.A., 2017. Lessons from #McKinney: Social media and the interactive construction of police brutality. The Journal of Social Media in Society 6, 284–313.

DeLisi, M., 2011. Criminal justice: Balancing crime control and due process. Kendall Hunt.

Engel, R.S., Corsaro, N., Isaza, G.T., McManus, H.D., 2020. Examining the impact of integrating communications, assessment, and tactics (icat) de-escalation training for the louisville metro police department: Initial findings. Alexandria, VA: International Association of Chiefs of Police. Retrieved 2, 194–202.

Fridell, L.A., Marier, C.J., 2023. The impact of suspect race and precipitating incident on community members' assessments of deadly force reasonableness. Homicide Studies 27, 142–166. doi:10.1177/10887679221112601.

Fridell, L.A., Richardson, D.A., Marier, C.J., 2023. Research in brief: The use-of-force "reasonableness divide". Police Chief Magazine URL: https://www.policechiefmagazine.org/research-in-brief-use-of-force-reasonableness-divide/.

Fulton-Babicke, H., 2018. "I can't breathe": Eric Garner and in/out-group rhetorics. Rhetoric Review 37, 434–446.

Gerber, M.M., Jackson, J., 2017. Justifying violence: Legitimacy, ideology and public support for police use of force. Psychology, Crime & Law 23, 79–95. doi:10.1080/1068316X.2016.1220556.

Jost, J.T., Stern, C., Rule, N.O., Sterling, J., 2017. The politics of fear: Is there an ideological asymmetry in existential motivation? Social cognition 35, 324–353.

Klein, G.C., 2018. On the death of Sandra Bland: A case of anger and indifference. Sage Open 8, 2158244018754936.

Klockars, C.B., 1985. The idea of police. SAGE Publications.

Kruschke, J.K., Liddell, T.M., 2018. The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. Psychonomic bulletin & review 25, 178–206.

Levitsky, S., Ziblatt, D., 2019. How democracies die. Crown.

Maher, G., 2021. A world without police: How strong communities make cops obsolete. Verso Books.

Marier, C.J., 2024. Justifiability and culpability in lethal self-defense: Police officers vs. civilians. Journal of Criminal Justice 90, 102142. doi:10.1016/j.jcrimjus.2023.102142.

McLean, K., Wolfe, S.E., Rojek, J., Alpert, G.P., Smith, M.R., 2020. Police officers as warriors or guardians: Empirical reality or intriguing rhetoric? Justice quarterly 37, 1096–1118.

Mourtgos, S.M., 2025. Probabilities Over p-Values: A Decision Framework for Evidence-Based Policing. CrimRxiv Https://www.crimrxiv.com/pub/3gbvojje.

Mourtgos, S.M., Adams, I.T., 2020. Assessing public perceptions of police use-of-force: Legal reasonableness and community standards. Justice Quarterly 37, 869–899. doi:10.1080/07418825.2019.1679864.

Mourtgos, S.M., Adams, I.T., McLean, K., Alpert, G.P., 2025. Risk and public judgments on police pursuits: A nationally representative conjoint experiment. CrimRxiv .

Mullinix, K.J., Leeper, T.J., Druckman, J.N., Freese, J., 2015. The generalizability of survey experiments. Journal of Experimental Political Science 2, 109–138.

Navarro, J.C., Hansen, M.A., 2023. "please don't be too nice": the role of political ideology in the approval of police use of force. Policing: An International Journal doi:10.1108/pijpsm-03-2023-0035.

Norris, R.J., Hicks, W.D., Mullinix, K.J., 2023. The politics of innocence: How wrongful convictions shape public opinion. NYU Press.

Patil, S.V., 2018. Public support for the punishment of police use of force errors: Evidence of ideological divergence and convergence. Police Quarterly 21, 358 – 386. doi:10.1177/1098611118766647.

Peters, B.G., Pierre, J., 2006. Governance, accountability and democratic legitimacy, in: Governance and Democracy. Routledge. chapter 7, pp. 115–132.

Pinker, S., 2012. The better angels of our nature: Why violence has declined. Penguin Publishing Group.

Richardson, D.A., 2022. The "reasonableness divide"": Comparing community members' assessments of force reasonableness to legal standards. Ph.D. thesis. University of South Florida. Tampa, FL.

Richardson, D.A., Fridell, L.A., 2025. The "reasonableness divide": Comparing community members' assessments of force reasonableness to legal standards. Justice Quarterly 42, 312–335. doi:10.1080/07418825.2024.2378712.

Saad, L., 2023. Historically low faith in U.S. institutions continues. URL: https://news.gallup.com/poll/508169/historically-low-faith-institutions-continues.aspx.

Schwartz, J.C., 2014. Police indemnification. NYUL Rev. 89, 885.

Silver, J.R., Pickett, J.T., 2015. Toward a better understanding of politicized policing attitudes: Conflicted conservatism and support for police use of force. Criminology 53, 650–676. doi:10.1111/1745-9125.12092.

Simon, C.A., Moltz, M.C., Lovrich, N.P., 2021. Police use of force and nativity: Revisiting standing evidence of public opinion on police use of force. Police Practice and Research 22, 1077–1094. doi:10.1080/15614263.2020.1772781.

Smith, D., 2020. Nine out of 10 Americans say racism and police brutality are problems, poll finds. The Guardian URL: https://www.theguardian.com/us-news/2020/jul/08/americans-racism-police-brutality-problems-poll.

Stinson, P., Wentzlof, C., 2019. Research brief one-sheet – No. 9: On-duty shootings: Police officers charged with murder or manslaughter, 2005-2019. Criminal Justice Faculty Publications URL: https://scholarworks.bgsu.edu/crim_just_pub/101.

Stoughton, S.W., Noble, J.J., Alpert, G.P., 2021. Evaluating police uses of force. NYU Press.

Strickler, R., Lawson, E., 2022. Racial conservatism, self-monitoring, and perceptions of police violence. Politics, Groups, and Identities 10, 254–275. doi:10.1080/21565503.2020.1782234.

Thompson, B.L., Lee, J.D., 2004. Who cares if police become violent? Explaining approval of police use of force using a national sample. Sociological Inquiry 74, 381–410. doi:10.1111/j.1475-682X.2004.00097.x.

Trahan, A., Russell, J., 2017. Race and police use of force: A regression analysis of varying situational approval from 1972 to 2012. Applied Psychology in Criminal Justice 13, 142–154.

USDOJ, 2021. Former Minneapolis Police Officer Derek Chauvin pleads guilty in federal court to depriving George Floyd and a minor victim of their constitutional rights. URL: https://www.justice.gov/archives/opa/pr/former-minneapolis-police-officer-derek-chauvin-pleads-guilt

Vitale, A.S., 2021. The end of policing. Verso Books.

Wilson, G., 1973. The psychology of conservatism (routledge revivals). Routledge.

Wolfe, S.E., Mclean, K., Alpert, G.P., Rojek, J., 2024. Us versus them? The problem of cognitive distortions in policing. Police Quarterly doi:10.1177/10986111241234310.