Document Title: Multi-Modal Analysis of Body-Worn Camera Recordings: Evaluating Novel Methods for Measuring Police Implementation of Procedural Justice

Author(s): Colby Dolly

Document Number: 311091

Date Received: January 2026

Award Number: 2020-R2-CX-0010

Multi-Modal Analysis of Body-Worn Camera Recordings: Evaluating Novel Methods for Measuring Police Implementation of Procedural Justice

Award: 2020-R2-CX-0010

Colby Dolly, Ph.D.
Director of Innovation and Science, National Policing Institute
cdolly@policinginstitute.org
National Policing Institute
2550 S Clark Street, Suite 1130
Arlington, VA, 22202
(202)-833-1460

September 16, 2020 to June 30, 2024

Award amount: $836,000

This project was supported by Award No. 2020-R2-CX-0010, awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of the Department of Justice.

TABLE OF CONTENTS

ABOUT THE AUTHORS

Colby Dolly, Ph.D.

Colby Dolly is a Director of Science & Innovation at the National Policing Institute. In this capacity, he is responsible for a portfolio of projects related to organizational assessments, policing improvements, policing accountability and reform, technology and innovations, and policing outcomes. Before joining NPI, Dr. Dolly held management roles for a Fortune 50 company and a large Midwest police department. In these roles, he was responsible for policy formation, data analysis, process improvement, and change management. Dr. Dolly holds a Ph.D. in Political Science from the University of Missouri St. Louis, where his research focused on organizational change and research methodology. In addition, he is a Prosci® Certified Change Practitioner and has a certificate in Lean Six Sigma.

Ellie Hansen, M.A.

Ellie Hansen is a Project Associate under the Science and Innovation team at the National Policing Institute. Before arriving at NPI, Ms. Hansen completed her Master of Arts program in Criminal Justice (MACJ) at Seattle University in the spring of 2022. As a graduate student, Ms. Hansen worked with the Office of Law Enforcement Oversight to provide research on police mental health responses and the efficacy of crisis intervention training across the country. She graduated from Christopher Newport University in 2020 after receiving a bachelor's degree in psychology and sociology with a concentration in criminology. Ms. Hansen's experience at NPI includes a diverse range of issues relating to policing, such as police culture, the organizational structure of agencies, critical incident reviews, technology and innovation, technical assistance and training, and officer safety and wellness.

Roberto Gallardo, Ph.D.

Roberto Gallardo has a doctorate in Sociology from the University of California at Riverside, specializing in criminology, race and class, and political economy. His research focuses on all aspects of policing in major cities. He has been recognized with grants and featured in publications like *The Criminal Justice Review* and *Race & Justice*. Currently, Dr. Gallardo is an Assistant Professor at the University of North

Texas at Dallas in the Department of Criminal Justice and Sociology, where he also serves as Program Coordinator. Dr. Gallardo also teaches courses for the Dallas Police Department.

John Coleman, Ph.D.

Dr. J. Eric Coleman is a Senior Lecturer in the Department of Criminal Justice and Sociology and the graduate program coordinator for the MSCJ degree program at the University of North Texas at Dallas. Dr. Coleman has over 30 years of traditional police experience, having served in various positions of police leadership. He is the former Associate Vice President of Campus Safety and Security and the University Police Chief at Texas A&M University-San Antonio. In addition, he served as Chairman of the Texas A&M University System's Police Chiefs Council and was a member of the Executive Board (Secretary) of the Texas Association of College and University Police Administrators. Dr. Coleman's teaching and research focus on police leadership, recruiting, police policy and management, and mental health in the administration of justice. In addition, Dr. Coleman holds a bachelor's degree in law enforcement and justice administration from Western Illinois University, a master's degree in political science and public administration from Midwestern State University, and a Doctorate from Nova Southeastern University in Higher Education Administration with a specialization in Criminal Justice Education.

Jonathan Wender, Ph.D.

Dr. Wender is co-founder and CEO at Polis. He is a "philosopher-cop" and entrepreneur who has dedicated his professional life to understanding and improving human interactions in high-consequence environments. He has broad experience developing technology, training, and technical assistance programs designed to improve public trust and safety. Dr. Wender leads the team that created the TrustStat™ BWC analytics system. He is a twenty-year police veteran and later served at DARPA as a senior advisor for the "Good Stranger" program. Dr. Wender also taught at the University of Washington. He is internationally recognized for his expertise on police decision-making, social interaction, and other topics. He earned his Ph.D. in criminology from Simon Fraser University.

Brian Lande, Ph.D.

Brian is a co-founder and Chief Science Officer at Polis. He is a fifteen-year police veteran and sociologist with a passion for using science and technology to improve public trust. He leads the practical application of scientific research on human interaction, trust, and policing in the TrustStat™ BWC analytics system. Dr. Lande also oversees Polis' research and evaluation portfolio on topics such as adaptive decision-making, de-escalation, and procedural justice. He directs liaison with Polis' university, non-profit, and other institutional research partners. Dr. Lande's pioneering work on computational social science led to his appointment as program manager at DARPA for the "Good Stranger" program. He holds a Ph.D. in Sociology from UC Berkeley.

Vivian Ta-Johnson, Ph.D.

Dr. Ta-Johnson is the Chief Data Scientist at Polis, where she directs the natural language processing functions of the TrustStat™ BWC analytics system and the related development of new AI tools for analyzing social interactions. She also serves on Polis research projects involving the application of natural language processing to the analysis of police-community interactions. Dr. Ta-Johnson is an Assistant Professor of Psychology at Lake Forest College in Lake Forest (IL), where she directs the Technology, Relationships, and Language (TRL) Lab. The TRL Lab focuses on using natural language processing and other computational techniques to analyze social interactions. Dr. Ta-Johnson holds a Ph.D. in experimental psychology from the University of Texas at Arlington.

Marc Tomlinson, Ph.D.

Dr. Tomlinson directs engineering and operations for the TrustStat™ system and supports related Polis projects on the use of AI to analyze high-consequence social interactions. Dr. Tomlinson has extensive experience on the modeling of psychological and social behavior using multimodal technology, deep-learning-based large-scale natural language processing, and leading-edge AI techniques. He has worked as a senior research scientist on several DARPA and IARPA programs on the application of computational social science, AI, and multimodal analysis to address complex defense and national security matters. Dr. Tomlinson holds a Ph.D. in cognitive psychology from the University of Texas (Austin).

Guy Ben-Yosef, Ph.D.

Dr. Guy Ben-Yosef is a Principal Scientist at GE Healthcare with interests in computer vision and Artificial Intelligence. He is a graduate of Ben-Gurion University of the Negev and has been a post-doctoral fellow at the Massachusetts Institute of Technology and Weizmann Institute of Science.

Peter Tu, Ph.D

Dr. Peter Tu Chief Scientist for Artificial Intelligence (AI) at GE Research. Since 1997, he has worked at GE Research on numerous DARPA, IARPA, DHS, FBI, and private sector projects on the application of AI, computer vision, and multimodal technology to the analysis of social interactions. Peter has published over 75 peer-reviewed papers and filed more than 50 patents. He is a graduate of Oxford University and the University of Waterloo.

SUMMARY OF THE PROJECT

The goals of this project were twofold: First, to develop novel multimodal techniques for automatic analysis of body-worn camera (BWC) recordings of police and community interactions to evaluate officers' adherence to procedural justice principles. Second, to validate the ratings generated by the automated process comparing software ratings of BWC videos to evaluations performed by human raters under high and low procedural justice conditions. The research questions below arise from this second objective. Both goals have been successfully met. This report will examine the methodologies employed, the findings obtained, and the implications of these achievements.

Research Questions

The project sought to answer the following research questions:

1. Are there differences in perceptions of procedural justice between (a) community members, (b) university faculty members and graduate students, and (c) police supervisors?

2. Do procedural justice scores generated by automated video analytics align with scores generated by (a) community members, (b) university faculty members and graduate students, and (c) police supervisors?

3. Can data from manually coded interactions be used to refine the automated coding algorithms and scoring/weighting procedures?

Research Design, Methodology, Analytical and Data Analysis Techniques

The methodology employed in this study was divided into three key phases. First, Polis Solutions began developing their automated software TrustStat, which aimed to systematically evaluate body-worn camera (BWC) footage for procedural justice metrics. The process began with a thorough review of existing coding instruments from structured social observation studies, particularly those by Mastrofski et al. (2015), to form the basis of the automated analysis tool. The software was designed to analyze verbal and non-verbal cues from BWC footage, identifying behaviors aligned with procedural justice principles. The second phase involved creating a coding instrument for human evaluators structured around the four dimensions of procedural justice: neutrality, participation, trust, and respect. Human coders, including community members, police personnel, and academic faculty, were trained to use this instrument to rate the videos. The third phase compared the TrustStat evaluations with human coders' evaluations to validate the tool's accuracy and reliability. This comparison tested whether the automated ratings aligned with human perceptions of procedural justice in police-community interactions.

Expected Applicability of the Research

The expected applicability of this project is significant for law enforcement agencies aiming to enhance accountability and transparency in police-community interactions. Using automated tools, such as TrustStat, to evaluate the use of procedural justice could provide law enforcement agencies the opportunity to assess officer performance in a new way. Such tools can serve as a robust mechanism for performance evaluation, moving beyond traditional reliance on subjective reports or sporadic incidents to a more data-driven approach. The insights gained from TrustStat can inform training programs, policy adjustments, and early warning systems, ultimately fostering improved community relations and trust in the police. Furthermore, the research validates the feasibility of using advanced AI-based analytics to handle large volumes of BWC footage, making it possible to assess officer interactions comprehensively and consistently. This transformation has the potential to reshape the way police performance is measured and managed, emphasizing the quality of interactions over mere output metrics like the number of arrests or tickets issued.

Participants and Collaborating Organizations

Carrying out this research project required the collaboration of several organizations, including the National Policing Institute (NPI), Polis Solutions, the Caruth Policing Institute from the University of North Texas at Dallas (CPI), and the Dallas Police Department (DPD). The NPI was the prime award recipient, with Polis Solutions and CPI providing support as sub-recipients. The NPI oversaw the project and carried out the evaluation of the tool created by Polis Solutions, with CPI providing support for coder selection and evaluation. The DPD served as an integral partner by providing videos for the creation of TrustStat, a multimodal software platform able to evaluate the use of procedural justice by police officers in the field.

Changes in Approach from Original Design

The project experienced delays and changes over its duration. The COVID-19 pandemic and the 2020 public demonstrations calling for police reform substantially delayed the project in the beginning as organizations were figuring out protocols for meetings and new working arrangements. These challenges required a necessary justification and revision of the original scope of work.

Additionally, the scope of the original study was adjusted to better reflect the current conditions in policing and police-community relations. The justifications for the original scope of work favored an alternate testing strategy over a field experiment to better inform the utility of the automated BWC classification tool. The original project proposed that officers at the DPD attend a 32-hour procedural justice training course. Given historically low staffing levels and constrained resources, this course would remove officers from the street for a considerable amount of time during their work week. Secondly, the DPD currently provides a substantial amount of procedural justice training to all officers, and it is unclear if additional training would yield any behavioral benefits.

Phase II of this study was redesigned to directly address the relationship between stakeholder perceptions of procedural justice and automated video scoring. This revision focused on rigorously testing and evaluating Polis Solutions' automated BWC analysis platform, TrustStat, and comparing it to how different stakeholders perceive elements of procedural justice in police-community interactions. In other words, the study applied sociological and empirical findings about how social judgments of the same stimulus vary in structured ways relative to stakeholders' social positions. TrustStat system accuracy would mean demonstrating the ability of AI technology to handle the empirical reality that the "ground truth" of social phenomena, such as respect, trust, participation, and so forth, vary in determinate ways (Fourcade, Lande, and Schofer 2016; Forgas and Bonds 1985;).

OUTCOMES

Statement of Problem

Compared to other professions, policing is uniquely difficult to study because what officers do in the course of a day is challenging to monitor and evaluate. Officers go from one assignment to the next, interacting with community members in very different circumstances, ranging from friendly encounters while getting coffee at a convenience store to having intense psychological confrontations when a patron refuses to leave a bar. Most of these interactions go undocumented with little follow-up on the quality. Contrast the duties of an officer with those of many public-facing jobs where work quality is easier to observe, and the consequences may not be as detrimental compared to policing. The oversight problem was identified decades ago and continues to present challenges as departments strive to optimize performance and hold officers accountable. BWCs, which emerged as a viable technology over a decade ago, present the potential to address this issue by providing insight into the interactions between officers and the community.

The use of BWCs by law enforcement agencies in the United States has proliferated since the 2010s. The current widespread use of BWCs makes it easy to forget how rare they were just a few years ago. A 2010 *TechBeat* story treated BWCs as a novel technology that is unknown except for a few innovative law enforcement agencies (National Law Enforcement and Corrections and Technology Center, 2010). In 2012, a National Institute of Justice (NIJ) report stated that "a major issue with the use of BWCs is a lack of technical standards and operational standards for protocols and procedures" (NIJ, 2012:missing page#?). However, despite the technical and regulatory challenges highlighted in the NIJ report, the use of BWCs continued to proliferate in the ensuing years. By 2016, 47% of law enforcement agencies had acquired BWCs, including 80% of large agencies (Hyland, 2018). By 2020, 62% of agencies were using BWCs, including over 90% of large police departments (Goodison & Brooks, Connor, 2023). In sum, what began less than twenty years ago as a niche technology of uncertain importance is now almost universally regarded as an essential requirement for any professional law enforcement agency.

Although the use of BWCs has expanded rapidly, the capacity to efficiently analyze the enormous amount of data collected by BWCs lags far behind. As a result, the broader potential of BWCs to improve practices and outcomes of policing has yet to be fully realized (White & Fradella, 2016; White et al., 2018). For the most part, law enforcement agencies have adopted BWCs out of political and legal necessity rather than treating them as a strategic data source to be systematically leveraged to support core agency and community objectives. In recent years, however, the idea that

BWCs are a strategic data source is gaining traction as more agencies and other stakeholders realize they were over-optimistic in earlier predictions that merely deploying BWCs on officers would positively change officer and public behavior. Lum et al. (2019) argue in their meta-analysis of studies on the efficacy of BWCs that the mere presence of cameras does not appear to influence either officer or community behavior significantly.

The rapid evolution of BWCs from novel technology to ubiquitous, even mandatory devices offers an important historical perspective for understanding that the policing profession now stands at the threshold of a new and closely related technological revolution: BWC data analytics. While it may take time to become apparent to the casual observer, the rapid adoption of BWCs has created a parallel massive growth in the amount of BWC data being recorded and stored. The proliferation of BWCs has created the world's largest and most valuable source of data on policing. According to one recent estimate, Axon's BWC data storage now exceeds 100 petabytes (PB) (Farooq, 2024). To put this perspective, when Axon initially moved its Evidence.com BWC data storage to Microsoft's Azure cloud in 2018, the transfer was only around 20PB (Dignan, 2019).

Both in terms of quantity and quality, BWC data is without comparison in its sheer magnitude and importance. Even the smallest police agencies can quickly amass thousands of hours of BWC video. The largest agencies record millions of videos per year. To take the example of the Dallas Police Department (DPD), the agency partner for the project described in this report, DPD patrol officers in 2023 recorded an estimated total of 200,000 videos per month, equivalent to nearly 80,000GB of data. While it may be easy to regard the massive volume of BWC video as a statistical abstraction, it is crucial to remember that BWC data has the untapped potential to humanize policing in ways that are just starting to become technically feasible at this dawning moment of BWC data analytics. Although some agencies are auditing footage, there is no published research on the volume or rigor of auditing across the nation and it likely varies considerable with the many agencies in the United States. Efforts should be made to standardize auditing methodologies to systematically check the quality of officers' performance for a wide range of interactions with the community.

At its root, policing comprises a vast tapestry of face-to-face interactions between officers and community members. Making practical sense of these interactions at scale is as important as it is difficult. For decades, society has sought to understand, regulate, and reform policing to best serve the public good. Until now, all efforts at improving and modernizing policing have been hampered by the profound opacity of what occurs when community members and officers interact. Law enforcement agencies, researchers, and a wide range of other government and private-sector organizations collect data on policing and seek to analyze it in valuable ways. However, efforts to understand, at scale, the street-level realities of policing have always been fundamentally limited by their dependence on relatively small samples of interactions.

Equally so, research on policing is skewed by an emphasis on incidents involving force or violence, which only account for approximately 1-3% of police-community interactions nationwide. Although 97-99% of police interactions involve no use of force, little is known about the complex factors that account for the remarkable capacity of police officers and community members to mitigate and entirely avoid conflict and violence. As a result, society lacks a tractable, data-driven understanding of why nearly all police-community interactions are non-violent. The result is a constant process of "reverse engineering," in which police professionals and researchers over-rely on trying to argue counterfactuals – how, for example, a given officer-involved shooting could have been prevented – as opposed to looking at large-scale data on non-violent interactions and modeling why peaceful outcomes are the norm rather than the exception.

The advent of BWCs has transformed the availability of data on police encounters. However, until recently, even when BWC video has become available, making practical use of it has been painstakingly expensive and time-consuming. For example, Federal Law Enforcement Training Centers (FLETC) BWC researcher Dr. Laura Zimmerman estimates that hand-coding a data set of approximately 500 BWC videos took more than two thousand hours, the equivalent of one full-time employee's annual work (Zimmerman, 2023). Moreover, the cataloging and sampling of BWC videos has been limited by the lack of technology that can automatically analyze, label, and sort large volumes of BWC data. The result is that most law enforcement agencies have no idea what is stored in their BWC video repositories, let alone have the capacity to analyze and use BWC data in practical ways.

In the last several years, researchers have begun to leverage BWC content as data for research and training. Makin et al.(2017, 2018) and Koslicki et al. (2019) have used BWC data to study predictors of the duration and level of police use of force (as well as contextual determinants (individual, situational, and environmental) of negative emotions in police-community encounters. Makin and his colleagues have also begun using increasingly sophisticated mathematical techniques to analyze BWC data. Others, such as Worden and McLean (2014, 2017), have used in-car dashboard camera content to code police and community member behavior in procedural justice constructs, such as giving voice, neutrality, quality of treatment, and trustworthy motives. Using data coded from video, Worden and McLean examined various factors (e.g., individual traits of officers and citizens, nature of the call, and levels of resistance) that shaped the level of procedurally just actions in a police-citizen interaction. Friis et al. (2020) analyzed BWC footage from transit fine enforcement officers in Denmark to identify the role that procedurally just actions play in mitigating the escalation of encounters.

Studies summarized in the preceding paragraph demonstrate that BWC footage can be used as a data source for police performance. However, for the video to be reviewed and turned into structured data, new methods will be required to make it feasible for large-scale research or applied use in law enforcement agencies. They further reflect an increased understanding that BWC data is about interactions, not just individuals. By sequentially coding data based on the "give and take" over time between officers and community members, these studies align with the theoretical and methodological work of Alpert and Dunham (2004) and Sykes and Brent (1983), who argue that if we want to understand police behavior, we must situate it in the temporal dynamics of reciprocal exchanges and transactions between participants (also see Collins, 2008; Goffman, 1967). There have been some insightful studies examining other critical aspects of the language used in police-community interactions, including research on respect (Voigt et al*.*, 2017), institutional dialogue (officers' requests, commands, questions and statements) (Prabhakaran et al., 2018), and features of procedural justice, such as how officers respond to community members' questions and requests (Mastrofski et al., 2015; Worden & McLean, 2014; 2017). While there was a benefit in these early efforts to systematically analyze BWC data and construct sequentially organized time-series datasets of changes during interactions, they have profound limitations.

First and foremost, prior research uses methods that rely on hand-coding large amounts of video, which is a laborious and inefficient process that cannot be scaled up. Depending on the measurement strategy, coding 1 to 10 minutes of video data can easily take an hour or longer. Second, the existing research faces significant issues of inter-rater reliability. Coders often have difficulty identifying and agreeing on the behavioral anchors on which they base their judgments on how to code a "slice" of behavior or even an entire video segment.

The multimodal BWC data analytics technology described in this report is the first of its kind and entirely unfamiliar to all but a few law enforcement agencies. However, suppose this history of BWCs is any indication of future events. In that case, it is safe to predict that BWC data analytics technology will rapidly proliferate in the next few years and will become as commonplace and vital to policing at BWCs themselves.

Literature review

Police Accountability and BWCs

The significance of video recordings of police behavior in the United States has been in the spotlight ever since the 1991 beating of Rodney King, which was captured on a camcorder by George Holliday, who lived near the scene. During the 81-second segment of the nine-minute video, King was kicked, shocked, and beaten with police batons. The video of the beating was given to KTLA, a local television station, and became national news within a few days of its release. Subsequent controversies around the police use of force have continued to be a dominant sociopolitical issue. In 2014, President Obama launched the President's Task Force on 21st Century Policing (PTFCP 2015) to deal with the crisis of police accountability (St. Louis et al., 2019). The high-profile deaths of Michael Brown in Ferguson, Eric Garner in New York, and Tamir Rice in Cleveland accelerated widespread demands for police officers to use body-worn cameras (USA White House, 2014; Ariel et al., 2014; Jameel & Bunn 2015). However, despite the hypothesized potential of body-worn cameras to improve police-community interactions, skeptical voices cautioned that BWCs could only address the consequences rather than the causes of poor police-community relations (e.g., Feeney 2015; Gonzales & Cochran, 2017).

Recently, George Mason University's Center for Evidence-Based Crime Policy launched a four-part research project that examined what needed to be known about using BWCs. The project found that out of 42 studies, the dominant theme measured complaints and officer use of force. Not many studies addressed the relationship between BWCs and the community attitudes and perceptions of law enforcement and their legitimacy (Crow et al., 2017).

Jennings et al. (2014) found in their study in Orlando, Florida, that police officers were generally supportive of using BWCs before the technology was introduced for use in the department. In the follow-up study, the researchers conducted a year later, three out of four officers required to wear BWCs thought that officers in the department should be required to wear the cameras. Most of the officers surveyed believed the cameras would help improve police tasks such as evidence collection, report writing, and other aspects of police work. The same survey participants from the Orlando Police Department were not convinced about the impact of BWCs on officers or community members, individuals, or public conduct (Jennings et al., 2015). Having more BWC experience is also associated with more positive views of the city-wide adoption of BWCs (Goetschel & Peha, 2017). Police leaders who were surveyed about the use of BWC were generally supportive of the use of cameras. Half of the police leaders surveyed believe that cameras would reduce unwarranted community member, individual, or public complaints and improve community member, individual, or public behavior in law enforcement contacts. One in five law enforcement leaders believe that the use of cameras would affect officers' behavior. They also felt that the media would use the videos of police encounters to embarrass the police (Smykla et al., 2016).

The public and policymakers have been seeking changes in how police are held accountable for their conduct in their official capacity. Many suggestions have been made to improve police accountability, but none have been more popular than implementing body-worn cameras (Miller, Toliver, and PERF, 2014). BWCs represent a viable means to monitor officers' day-to-day activities and make judgments on performance without relying on officer and citizen accounts alone.

Perceptions of Police

Public perceptions of law enforcement are a complex amalgam founded on individual experiences, cultural narratives, and historical realities. These personal and often emotionally charged perceptions can profoundly impact the relationship between law enforcement and the communities they serve. This section delves into the intricate landscape of public perceptions of the police, examining factors such as variations of perceptions based on critical demographic variables rooted in individual encounters with law enforcement, historical context, and social and economic realities. By acknowledging the complexities and nuances of public perceptions, we can foster a more equitable and collaborative approach to community safety and justice.

Early 2000s research by Benedict et al. (2000) found that many of the public generally held positive views of law enforcement. These positive perceptions are likely attributed to feelings of safety and security provided by law enforcement or even positive personal interactions with individual officers. However, Worrall (1999) rightly points out a key factor lacking in studies: the lack of consideration for various demographic variables. Many early studies on police perceptions treated the public as a homogenous group; this approach masked significant underlying differences. A clearer picture emerged when researchers began to disaggregate the data by factors such as race, ethnicity, socioeconomic status, and other significant demographic variables. This disaggregation revealed not only a more nuanced understanding of public sentiment but also the stark and essential differences in the perception of law enforcement by different groups. Local perceptions of the police also differ from global perceptions of the police, with local perceptions tending to be more positive compared to global perceptions (Griffith & Foley, 2020; Perkins, 2016). Fueled by many historical and social factors, these differences have become central to understanding the complexities of the relationship between the police and the public.

In their attempt at a meta-analysis, examining over 100 articles on attitudes toward and perceptions of the police, Brown and Benedict (2002) build on earlier research and discover that four variables—age, neighborhood, race, and contact with the police—have been found to influence attitudes and perceptions toward the police consistently. More recent studies have examined the relationship between race and the perception of the police and have come to similar findings about the influence of race on the perception of policing. Expanding on the work of Brown and Bowen (2002) in a more recent meta-analysis of perceptions of police, Peck (2015) provides a comprehensive literature review of empirical studies on perceptions and attitudes towards the police across various racial and ethnic groups. Her finding on the examination of 92 articles confirms the consistent pattern: individuals who identify from minority populations are more likely to hold negative views and attitudes towards law enforcement in comparison to Whites. Specifically, Black males have the most negative perception of law enforcement. A more recent experimental study of how different social groups judge the procedural justness of simulated traffic stop videos also showed that there are differences in the baseline ratings between social groups (Johnson et al., 2017). Although procedurally just actions improved assessments of the police, the scores were more negative for Black respondents compared to other groups. Nadal et al. (2017) also found that among men, Black men were more likely to report negative perceptions of police compared to White and Asian men. As it relates to other ethnic groups, Hispanics hold a nuanced position. Some studies suggest Hispanics may have more positive views of law enforcement compared to Blacks (Schuck & Rosenbaum, 2005). Other research also indicates they perceive the police more negatively than Whites (Huggins, 2012; Peck, 2015). These differences in perspective highlight the need for further research to understand the specific experiences and perceptions of Latino/Hispanic communities. For example, Barboza (2012) found that Mexican-Americans with a more substantial group consciousness are more likely to report negative perceptions of treatment by law enforcement.

Additionally, Asian Americans also present a complex picture when it comes to perceptions of the police. Some studies find that Asian Americans generally hold more positive perspectives of law enforcement than other non-white groups (Callanan and Rosenberg, 2011; Wu, 2014). However, the positivity among Asian-American groups is not absolute, similar to Hispanics; complexity stems from the heterogeneity within the Asian-American population. Asian Americans encompass diverse ethnicities and nationalities, each with potentially distinct experiences and cultural perspectives impacting their perception of law enforcement.

Among other demographic variables examined that potentially impact the perception of the police include gender and age. Findings regarding gender and perception of the police are less conclusive. Some studies suggest females may have more positive perceptions of the police, while others find no significant difference (Brown & Benedict, 2002; James & Lee, 2015). This inconsistency necessitates further research to understand the nuances of gender and its influence on perceptions (Nedal et al., 2017). Research also suggests a potential correlation between age and perceptions, particularly among young people of color. Rengifo et al. (2015) highlight that young people of color report experiencing more negative interactions with the police compared to older adults within their racial groups.

Research also suggests that social class and neighborhood context influence perceptions of police (Schuck et al., 2008). For example, Schuck et al. (2008) found that middle-class African Americans and Hispanics living in disadvantaged neighborhoods reported increased unfavorable attitudes toward law enforcement than those in more advantaged areas. Similarly, Wu et al. (2009) showed that African Americans in economically advantaged neighborhoods were less likely to be satisfied with police than whites in the same neighborhoods. As described, neighborhood characteristics also play a role in shaping attitudes toward the police. Studies by Reisig and Parks (2003) and Sprott and Doob (2009) highlight that residents in neighborhoods with concentrated poverty and high crime rates tend to have lower levels of satisfaction with police. Conversely, alternative patrol strategies like foot patrols and positive police behavior are associated with higher satisfaction (Reisig & Parks, 2003). Payne and Gainey (2007) further emphasize how perceptions of safety can influence attitudes toward law enforcement, finding that residents who report feeling unsafe or who are approached by drug dealers have more negative views of the police.

Schuck and Rosenbaum's (2005) research adds another layer by differentiating between global and neighborhood attitudes toward the police. Their findings suggest that negative personal encounters with law enforcement can negatively impact both a resident's overall perception of police (global) and their perception of police within their neighborhood (Schuck & Rosenbaum, 2005). Interestingly, positive non-enforcement contacts within a neighborhood can improve residents' perception of police, specifically in that area. In several studies exploring the link between police-resident interactions and public perception of police, positive interactions with officers generally improve resident satisfaction (Wentz & Schlimgen, 2012; Peyton et al., 2019). However, in general, studies find that more negative interactions with the police are more impactful in shaping perceptions of the police than positive interactions (Skogan, 2006).

Furthermore, Schuck and Rosenbaum’s (2005) study found a stronger connection between global and neighborhood attitudes for African Americans and Latinos compared to Whites. These findings align with research by Lai & Zhao (2010) and Schuck et al. (2008), suggesting that negative police interactions have a more significant impact on minority communities' perception of police in both general and local contexts (Schuck & Rosenbaum 2005; Lai & Zhao, 2010; Schuck et al., 2008). Additionally, Livingston et al. (2014) emphasize the importance of fair and respectful treatment in all police interactions. Their research shows that even people with mental illness report satisfaction with police interactions when treated fairly. According to Wells (2007), the most crucial element influencing people's opinions of police during interactions is treating them fairly.

National outrage and widespread protests under the banner of "Black Lives Matter" thrust the issue of race and policing into the national spotlight. Research suggests a potential amplification of negative perceptions towards the police, especially among Black Americans, following the various high-profile cases involving the death of citizens at the hands of the police. Specifically, these events have promoted increased distrust and skepticism, reduced satisfaction and faith in policing, and decreased overall legitimacy (President's Task Force on 21st Century Policing, 2015). Most research indicates that positive contact with the police improves perceptions of the police, while negative contact has the opposite effect (Worrall, 1999). Nevertheless, research has begun to increase our understanding of police and community in the current political and social climate (Skaggs et al., 2022; Morrow et al., 2021; White & Ferrandino, 2022), the long-term consequences of these incidents on the public's perceptions are still unfolding, it has undoubtedly contributed to a heightened awareness of the complexities surrounding police-community relations.

Public perceptions of law enforcement hold significant weight for several reasons. Firstly, they directly impact public trust and cooperation. When people trust the police, they are more likely to engage in behaviors that assist law enforcement, such as reporting crimes and cooperating with investigations. Conversely, negative perceptions can breed fear, distrust, and a reluctance to interact with the police, hindering their ability to protect the community effectively. Secondly, perceptions influence the legitimacy and accountability of the police. Positive perceptions contribute to the view of law enforcement as legitimate and fair, strengthening their social contract with the community. Conversely, negative perceptions raise concerns about police legitimacy and accountability, potentially leading to social unrest and a loss of faith in the institution. Reductions in police legitimacy and increased anti-police sentiment may have even diminished individuals' interest in pursuing a career in law enforcement (Copeland et al., 2022), especially among college students (Morrow et al., 2021).

Thirdly, perceptions impact effective policing. Positive perceptions can foster positive interactions between law enforcement and the community, leading to better information-sharing, problem-solving, and community-oriented policing strategies. Conversely, negative perceptions can create barriers to communication and collaboration, hindering effective policing efforts and potentially escalating tensions. Finally, understanding public perceptions is crucial for informing policy and reforms within law enforcement agencies. By understanding public opinion factors, policymakers and police leaders can develop strategies to address concerns, build trust, and improve police-community relations. In essence, perceptions of the police significantly impact various aspects of community safety, law enforcement effectiveness, and the relationship between the police and the public. Understanding and addressing these perceptions is crucial for building trust, promoting cooperation, and ensuring a fair and effective criminal justice system.

Procedural Justice

Evaluating police performance while interacting with citizens is challenging. At a basic level, police supervisors often use complaints, or the lack thereof, as the basis for judging the quality of officers' performance. Relying solely on complaints does not capture the full scope of police-community interactions. Of course, police supervisors could randomly watch officers on calls interacting with the community, but this, too, has problems of objectivity because the presence of the supervisor is likely to change behavior. A promising paradigm to understand police performance as it relates to direct interaction with the community is procedural justice. Popularized by Tom Tyler and expanded upon by other scholars (Mazerolle et al., 2013; Sunshine & Tyler, 2003, 2003; Tyler, 2006; Tyler & Huo, 2002), procedural justice posits that by involving community members in a fair and transparent process, they are more likely to accept outcomes that are contrary to their self-interest (Tyler & Huo, 2002, p. 7). Whether one interaction or thousands, using fair processes should lead to more collaboration and greater police legitimacy, according to the theory.

Incorporating procedural justice into police-community interactions offers substantial benefits. At the interaction level, where officers are face-to-face with community members, procedural justice has the immediate benefit of encouraging voluntary compliance with the justice system despite the potential for adverse outcomes for the individual community member involved. One must look no further than a routine arrest where the adverse individual outcome is obvious, but the fair and just behavior of the officers encourages compliance. At a more macro level, the systematic implementation of procedural justice may encourage communities to have greater faith in the criminal justice system to keep the peace and enforce laws, which, in turn, should lead to greater cooperation for crime control efforts.

The literature examined covers a wide range of topics but brings together a key set of concepts that are important to the community and the police. The concept or idea is that given the long-standing need for police oversight and accountability, BWCs should be used not only as a purely reactive mechanism for complaint adjudication but as a way to proactively monitor and shape police performance by systematically evaluating police-community relations. Procedural justice can serve as the theoretical frame by which to evaluate interactions. Given the varying perceptions of the police, it is critical to ensure that whoever or however the evaluations are done, they are aligned with community standards about police performance. Research looking at how different groups rate body-worn cameras (e.g., Johnson et al., 2017) suggests that different groups assess procedural justice differently or at least from different baselines, even though different groups will generally show responsiveness to the presence of procedural justice (e.g., favorability increases; Boivin et al., 2020). Other research shows that the baseline person perceptions of police officers differ between groups as well (Sim et al., 2020). For example, people's evaluations of police procedural justice are also shaped by their experiences with individuals and situations outside of the context of policing (Pickett & Nix, 2018). Nonetheless, there is reason to believe there is some invariance in judgments of procedural justice as ratings appear to improve with procedural justice irrespective of the central tendencies of different social groups (Wolfe et al., 2016). These concepts are the basis for the creation of TrustStat BWC analysis technology and the research into its applicability to evaluate BWC footage.

ACTIVITIES AND ACCOMPLISHMENTS

Analytical Plan and Methodology

The project team aimed to develop AI software that would systematically evaluate BWC footage for officers' use of procedural justice and compare the results to evaluations by human coders. To achieve these objectives, the project was divided into three phases, which included:

1. Develop the TrustStat AI software to evaluate BWC footage
2. Develop a coding instrument for human coders and assess coding consistency across human coders.
3. Compare the evaluation results of TrustStat and the human coders

The following sections will discuss the work completed in each phase, including the theoretical background and practical context.

Developing TrustStat

The Polis team began developing measures of procedural justice by reviewing coding instruments created and used across several structured social observation (SSO) studies. After reviewing the extensive literature on procedural justice, we found that the work derived from Jonathan-Zamir et al. (2015) was most relevant, as it focused on validating a coding instrument that could be used to structure observations of publicly visible human interactions. Structured social observation is a standard observational method that uses random sampling of police officers and pre-established coding systems to narrow the focus of observers, thus generating reliable and quantifiable data while minimizing observer bias (Worden & McLean 2014). One of the goals of Jonathan-Zamir et al. (2015) initial work was to systematize a way of observing police behavior across different times and research locations. Several other research studies of procedural justice have used similar instruments, which has allowed the instrument to continue being validated and reliably used across different research locations years apart (e.g., see Terpstra & Van Wijk, 2021; Worden & McLean, 2014 McCluskey et al. 2023; Mastrofski et al., 2015 McCluskey & Resig, 2017).

To develop TrustStat's procedural justice analytics, we needed a coding framework that had been applied to video data. Fortunately, several studies have applied the procedural justice coding framework developed by Jonathan-Zamir et al. (2015) to naturally occurring video-based data from routine police-citizen interactions, including in-car (dash) cameras (Worden & McLean, 2017) and body-worn cameras (BWC) (McCluskey & Reisig, 2017, McCluskey et al., 2023; also see a similar but simplified approach in Sytsma et al., 2021; Piza & Sytsma, 2022). These studies have demonstrated that the coding instrument works with video-based data and produces results like direct field observation by researchers. Prior research has further shown the benefit of using video as the data source for analyzing human interactions and social dynamics because it is structured, repeatable, transparent, and more accurate (Piza & Sytsma, 2023). Potential errors attributable to observer bias, inattention, memory lapses, and other factors can be mitigated (Terrill et al., 2023). However, as Terpstra and colleagues (2023) noted, this kind of study is still very labor intensive, an area where automated analysis can significantly impact.

A validated coding instrument of observable behaviors associated with procedural justice is important because the TrustStat system only “knows” the world through the sensory modalities available to it: the images and sounds captured by video. The TrustStat system is designed to “see” and “hear” empirically observable available features of social interaction. In part, the Polis team wanted to avoid trying to “mind-read” the participants in observed interactions, which would require the system to try to access “hidden” psychological states of participants. The team also thought such an approach would be theoretically misguided since social interactions are coordinated among individuals who form representations, plans, expectations, and understanding of their interactions based on their perception of the dynamic interchange of movements, bodily articulations, and sounds contained in social encounters. In other words, the team believed it was essential to pay attention to the dynamic details of the interaction and describe these features above and beyond the subjective states of the participants.

Consequently, our first step was to go through the various models and techniques available to us in the field of Natural Language Processing to see what tools could detect key features of verbal behaviors, semantics, affect, and so forth, which could be mapped onto the coding instrument developed initially by Jonathan-Zamir et al. (2015): 1) participation, 2) neutrality, 3) dignity and respect and 4) the apparent trustworthiness of the participants. Each of these dimensions of procedural justice are distinctive orthogonal aspects of social life, and each has unique challenges in accurately estimating the category. In the following section, we describe how we elaborated on Jonathan-Zamir's (2015) framework for the four dimensions of procedural justice.

Initial Dataset Description

Based on Jonathan-Zamir et al.'s (2015) work on behaviors that are consistent with procedural justice, the Polis team identified natural language markers that are representative of procedurally just behaviors that could be used to evaluate an officer's performance on procedural justice via BWC videos. In other words, the degree to which the natural language markers (i.e., procedurally-just behaviors) are present in a BWC video would indicate the extent to which the officer behaved procedurally-just in the given police-community encounter. Up to 24 natural language markers were selected to measure behaviors relevant to and representative of each of the four pillars of procedural justice. We also focused on selecting natural language markers that directly convey information about the speaker's thoughts, feelings, and intentions and can be accurately and reliably captured in BWC videos. Table 1 reports the natural language markers selected for each pillar.

Table 1. Natural Language Markers

Procedural Justice Pillar	Example Behavior	Example Natural Language Markers
Neutrality	Speaking in a manner that is easy to understand	acknowledgments, causation, first-person pronouns, first names, formal/informal titles, for me, for you, function words, gratitude, giving agency, hedges, lexical diversity, negations, number of questions asked, readability, reasoning, second-person pronouns, tentativeness, word count
Participation	Expressing interest in community member input	acknowledgments, asking for agency, first names, formal/informal titles, first-person pronouns, filler pauses, for me, for you, giving agency, greetings, number of questions asked, second-person pronouns, tentativeness, word count
Respect	Showing respectful behaviors	asking for agency, apologizing, formal/informal titles, gratitude, greetings, introductions, not using swear words
Trust	Showing care/concern	acknowledgments, apologizing, dominance, for you, giving agency, hedges, gratitude, greetings, negations, number of questions asked, reassurance, reasoning, safety, second person pronouns

To validate the structure of our language-based measure of procedural justice, 100 BWC videos were professionally transcribed, providing the verbal behaviors of all individuals in each BWC video. The camera-wearing officer’s verbal behaviors were analyzed by calculating the degree to which the natural language markers were present in each BWC video. This dataset was then submitted through a confirmatory factor analysis (CFA). The natural language markers loaded on four factors consistent with our conceptualization of the four pillars of procedural justice.

Technical Research Annotation, Population, and Validation

A team of trained student raters performed initial annotations (ratings) of BWC videos. Based on previous procedural justice research, an initial set of dimensions was developed for annotation, which included behaviors and features representing each of the four pillars of procedural justice, as well as an overall assessment of officers' performance on procedural justice (below average, average, excellent). The dimensions were revised and refined for optimal annotation based on rater feedback and analysis. After completing this process, all raters annotated a subset of BWC videos on the final set of dimensions. Table 2 provides examples of annotated dimensions for each pillar. All annotated dimensions can be found in Artifact 5. Raters were instructed to score each dimension based on the context of each video encounter. For instance, a primary police officer interrupting a primary community member once time during an encounter would not automatically yield a specific score on this dimension. Rather, the score on this dimension would depend on the contextual characteristics of that specific encounter. Annotation scores of dimensions nested under a given pillar were aggregated, resulting in four separate scores representing each of the four pillars of procedural justice. Analyses revealed that the aggregated scores for each pillar were significantly correlated with overall assessments of the officer's performance on procedural justice.

Table 2. Examples of Annotated Dimensions

Participation	Neutrality	Respect	Trust
PO addressed PCM's questions; PO provided information/ viewpoint/relevant facts for PCM.	PO explained the goal of interaction/why PO chose to resolve the situation as they did.	PO behaved/spoke respectfully to PCM during the encounter.	PO offered advice, assistance, and resources for PCM.
Frequency/Duration (Never, Very little, Some of the time, Most of the time, All of the time)	Was this evident in the encounter? (Yes, No)	Frequency/Duration (Never, Very little, Some of the time, Most of the time, All of the time)	Was this evident in the encounter? (Yes, No)
PO allowed PCM to speak/provide info without interrupting.	PO communicated with PCM while carrying out a task.	PO behaved/spoke disrespectfully to PCM during the encounter.	PO behaved/spoke in a casual/friendly way that helped build rapport and reduce tension.
Frequency/Duration (Never, Very little, Some of the time, Most of the time, All of the time)	Was this evident in the encounter? (Yes, No)	Frequency/Duration (Never, Very little, Some of the time, Most of the time, All of the time)	Frequency/Duration (Never, Very little, Some of the time, Most of the time, All of the time)
Note. PO = primary officer; PCM = primary community member			

Multimodal Analysis of Procedural Justice

The TrustStat system utilizes advanced AI and machine-learning techniques to automate the social analysis of police body-worn cameras. The video recordings from the camera are played through three separate classification components, which process the different modalities (visual, audio, and speech) present within the video. The visual information from the video is automatically analyzed by deep neural networks that detect the facial expressions (smiling, frowning, etc.) of the individuals visible to the camera. These networks have been trained on multiple large datasets covering a range of face types, ethnicities, and lighting conditions. In addition to facial expressions, TrustStat also utilizes additional deep neural networks to identify physical actions by individuals (e.g., someone sitting down, handcuffing, walking away, holding a weapon). These action sets are currently built and evaluated by request. It is important to note that TrustStat does not use facial recognition technology. The computer vision AI in TrustStat assesses facial expressions in order to analyze the nature and intensity of emotions, but it cannot identify specific individuals.

The audio information from the video is first analyzed to detect voices; these sections of the audio are then processed to identify the tone being carried within those voices. Tone covers both emotional signatures (anger, happiness) and vocal characteristics that might indicate pain or excitement. In addition, a speech-to-text engine is used to transcribe the audio and mark words spoken by the same person.

The transcribed audio is automatically processed for various features related to its social meaning. TrustStat utilizes a variety of pre-trained classifiers that identify if a particular utterance by an individual is a question, statement, or command; the affective content of the utterance, its valence (positive or negative) and intensity; its social implicatures (establishing credibility, challenging credibility, showing collegiality); and how the verbal interactions indicate the roles (officer/community member) and relationships between the speakers. As with the visual and audio classifiers, these primarily use a variety of neural networks that have been pre-trained on various open-source datasets covering various topics and often-time languages.

Natural Language Processing (NLP)

The transcript of a given BWC video is first restructured, and its contents are pre-processed to ensure that it is in an appropriate format for NLP analyses. Then, each speaking turn within the transcript is scored on the natural language markers pertinent to procedural justice. Resulting scores indicate the degree to which a given natural language marker was present in the corresponding speaking turn. Each natural language marker is detected using existing models of natural language, which are trained on standard American English. This means that all verbal behaviors—including local colloquialisms—are evaluated from a standard American English perspective. Accurately accounting for local colloquialisms requires all individuals in the BWC videos to use and understand local colloquialisms in the same manner. This is not possible to account for, especially as variations in the use and meaning of local colloquialisms can be found even among individuals from the same group/region. Next, the speaking turns belonging to the primary officer are identified using a language model developed to automatically identify the primary officer in a given BWC video with high accuracy. The primary officer's natural language marker scores are then aggregated to produce an overall procedural justice score and overall scores for each of the four pillars.

The data from the three sources are combined utilizing a log-linear model to produce final classifications for each of the four pillars of procedural justice (Respect, Participation, Neutrality, Trust) and an overall procedural justice rating for the officers, as well as to identify the level of respect, and disrespect showed by the community members and their level of cooperation with the officer. These classifiers have been specially trained on annotations of police/community interactions by academics and current or former police officers.

Diarization: Police Officers and Coding

Several challenges arose during the creation, training, and evaluation of the TrustStat software. The most significant potential complication is related to diarization. Diarization refers to the ability to accurately attribute language to a specific speaker. For example, in the analysis of BWC video, diarization enables the distinction between what an officer said and what a community member said. The more people who are present in a BWC video, the greater the challenge of diarization. Accurate diarization is also more difficult when there is significant background from sources such as traffic and road noise. The system does not use facial recognition to identify entities or speakers, and consequently, the system cannot use face detection to improve diarization by identifying who visually appears to be speaking.\ Diarization only marks sections of words spoken by the same speaker and similar facial expressions exhibited by the same individual within a video. TrustStat does not attempt to identify speakers across videos, nor does it preserve any features that would support this. Diarization is still an open problem, but due to the variety of sources of data (video, audio, transcripts), mistakes in diarization can be corrected to provide a reliable rating for the *interaction,* even if a particular utterance is misattributed.

The second complication that arose during the development of TrustStat was recruiting supervisors for development-related annotation. Police officers are overworked, and finding time to make annotations on videos is a challenge. However, as evidenced in the findings, ratings by community members and academics appear reasonably consistent with those of officers.

Student Coding

The main complication from the student annotation process revolved around time and effort. Raters had to complete annotation training before officially annotating BWC videos. During annotation training, raters became familiar with the audio and video formatting of BWC videos, the procedural aspects of police-community member encounters, the dimensions they would annotate, and identifying the primary community members. All raters were also given a small training set of BWC videos to practice annotation. These videos were used for training purposes only. As raters annotated the practice videos, their progress was continuously evaluated to ensure comprehension on all levels of the annotation process. They were also instructed to seek guidance and clarification on any aspects of the annotation process that they found confusing or unclear. Annotation training was complete once all raters could annotate BWC training videos without any additional inquiries. After they had completed annotation training, annotating one BWC video took a fair amount of time, given that raters needed to annotate over 50 dimensions for every BWC video. This was further exacerbated by longer BWC videos. Raters reported that the work required to annotate a BWC video was cognitively taxing. To minimize fatigue effects, raters could only annotate a limited amount of BWC videos daily, which lengthened the annotation process even more.

The Final Product

The developed platform, TrustStat, is a web-based AI software system for receiving and processing Body Worn Camera (BWC) footage from various sources. TrustStat is based on Microsoft's Azure Government Cloud. The system is multi-modal, analyzing audio and video from police BWCs to evaluate a wide range of social interactions, including procedural justice, de-escalation, compliance, conflict, detentions, and event classifications. It is highly scalable and capable of efficiently analyzing thousands of videos. Users can view uploaded videos and access basic summary information, with the ability to filter by individual officers, shifts, events, etc. When fully mature, TrustStat could serve many functions in helping police organizations understand the performance of officers in the field.

Develop a Coding Instrument for Human Coders

Similar to the development of the TrustStat coding algorithm, the development of the instrument for human coders was based on previous research using systematic social observation of police activity (Jonathan-Zamir et al., 2015; Worden & McLean, 2017). Also, like the instrument used for TrustStat and following the literature, the instrument was organized around four dimensions of procedural justice: neutrality, participation, trust, and respect. Unlike previous instruments found in the literature, such as the one used by Jonathan-Zamir et al. (2015), the instrument used for this project was not based on dichotomous decision points where points were allocated based on an officer exhibiting behavior associated with procedural justice, such as "exhibiting intermittent respect." (Jonathan-Zamir et al., 2015). Instead, each item's responses were based on a seven-point Likert scale where the respondents selected from strongly disagree (1) to strongly agree (7). The coders were instructed to consider the actions of the officer throughout the duration of the video when making a selection on the scale.

The rationale for using a Likert scale to rate the entire video is that police-community interactions are dynamic, and officers may use greater or lesser amounts of procedural justice depending on a multitude of factors that are specific to the particular point during the interaction. For example, an officer may use very little procedural justice at the beginning of an interaction or act impolitely only to reverse course and end the interaction by providing helpful guidance to a community member. Because of the highly dynamic nature of the interactions, the research team believed that a Likert scale rating of the entire video would do a better job of capturing the intricacies of the interactions. Additionally, this type of rating system would better align with the manner in which TrustStat rates videos, looking at every instance of the interaction and creating a collective score.

The coding instrument found as Artifact 1 used 35 items or questions for each video. Eighteen of the items related directly to the dimensions of procedural justice. Of those 18 items, five were to rate neutrality, three for participation, seven for trust, and four for respect. Other items on the survey were variables used to collect demographic and group information from the respondents. The instrument was delivered to the coders through Qualtrics, where the coders viewed each video and completed the 35 items. Before using the instrument, NPI staff tested the instrument to help determine the amount of training instruction needed to use it for untrained coders. Based on this testing, the coder instruction session was created.

Once the test videos and coders comprised of community members, faculty and graduate students, and police supervisors were in place, the project team reserved a computer lab at the University of North Texas at Dallas, where the coders could view the videos and rate them using the instrument created for the project. Before coding the videos, coders checked in with project staff, and then coders were provided with all ten test video files along with a link to the coding instrument. The coding was conducted in person because of the complex nature of the instrument and the videos. The project team wanted to ensure the coders were provided support if they had questions or needed clarification. Additionally, having the coding done in a supervised setting minimized the risk of the test videos being recorded or sent directly to individuals outside the research project.

Selection of Coders

For Phase II of the study, UNTD faculty recruited coders from three groups: DPD supervisors, faculty members, graduate students, and community members. For police supervisors, UNTD faculty coordinated with the DPD to identify potential first-line supervisors as participants. They reached out to existing DPD administrator contacts via phone. Once officers were identified, researchers contacted them by email. The email script can be found in Artifact 2.

To recruit UNTD faculty members and graduate students, the research team worked with program coordinators to identify participants. Faculty members from social science disciplines were prioritized, and if sufficient faculty members were not recruited, the research team recruited social science graduate students at UNTD, similar to the faculty researchers. Although the faculty members and graduate students possessed knowledge of social science, none of them were particularly knowledgeable about procedural justice. Potential participants were recruited by a standardized script found in Artifact 2.

Community member participants were recruited by the SERCH (Service, Education, Research, Community, and Hope) Institute at UNTD. SERCH was used to leverage its contacts within the local Dallas community. SERCH sought to identify participants from diverse backgrounds for the research but did not use a systematic method to select the participants. Table 3 shows the key characteristics of all respondents participating in the coding (based on self-reporting). The number of minority respondents is high across all groups, and as expected, the education levels of faculty members are the highest, with 82.6 percent having a master's degree or higher.

Table 3. Demographics of Community Member Respondents

Characteristic	n	%	n	%	n	%
	DPD Supervisors		Faculty and Graduate Students		Community Members	
Gender						
Male	8	61.5	11	47.8	11	36.7
Female	5	38.5	10	43.5	19	63.3
N/A	0	0.0	1	4.4	0	0.0
Race						
Non-Hispanic/White	2	15.4	11	47.8	1	3.3
Hispanic/Latino	4	30.8	3	13.0	7	23.3
Black/African American	7	53.8	4	17.4	19	63.3
Asian	0	0.0	4	17.4	1	3.3
Pacific Islander	0	0.0	0	0.0	2	6.7
N/A	0	0.0	1	4.4	0	0.0
Education Level						
High School Graduate	0	0.0	0	0.0	2	6.7
Some College (No Degree)	2	15.4	1	4.4	9	30.0
2 Year Degree	1	7.7	1	4.4	8	26.7
4 Year Degree	7	53.8	1	4.4	8	26.7
Master’s Degree or Higher	3	23.1	19	82.6	3	10.0
N/A	0	0.0	1	4.4	0	0.0
Mean Age	45.6		47.6		33.1	

Once community stakeholders were identified, researchers contacted them via email using a similar script used with police supervisors. The full script can be found in Artifact 2. A total of 66 participants from February to May of 2024 completed coding sessions (community members: n=30, police supervisors: n=13, faculty members and graduate students: n=23).

Selection of Body Worn Camera Video Recordings

The videos used for the research were provided by the DPD. The DPD excluded videos that were prohibited from viewing because of ongoing investigations or evidentiary value. Although the videos used for this research have no evidentiary value, they would typically not be released to the public unless there was a request via a Freedom of Information Act or similar request under State of Texas law. Videos of police-community interactions were consistent with the details described in Table 3 below. In general, videos were selected to capture an interaction between a single officer and a single community member. Videos were determined to be excluded if the audio was not clear or in any language other than English. The final test sample did not include speech that was inaudible or recordings in noisy environments. Additionally, to reduce risk to project participants, videos were excluded if they involved on-video violence or substantial injury.

During the video review, the project team identified videos with some diversity in the tone or nature of the interaction. The selection process sought to identify an equal number of videos where the officer uses higher and lower levels of procedural justice. For example, the team sought to identify videos where officers are polite and allow the community members to explain themselves. Conversely, the team attempted to select videos where officers seemed uninterested and interrupted the community members.

For this study, recordings were to capture police contact with community members in a wide range of settings and contexts, such as calls for service, investigative stops, traffic stops, contacts with suspicious persons, and contacts with persons involved in minor offenses (drinking in public, loitering, etc.). The project team used complete, unredacted body-worn camera footage that recorded the interactions from the moment the officer arrived on the scene to the completion of the call.

This study utilized a sample of 100 unredacted body-worn camera recordings from the DPD. The initial sample of recordings was electronically and securely transferred to NPI. The recordings were entirely separate from a larger corpus of videos that had been used to support training of the TrustStat software. Recordings were reviewed and classified into exclusion and inclusion categories set by IRB guidelines. The following two sections will discuss the details of inclusion and exclusion criteria to determine the final sample of test videos. The inclusion and exclusion guidelines are found in Table 4. From the initial sample of 100 test videos, ten videos were selected with police-citizen interactions ranging from low, neutral, and high perceptions of procedural justice. Several tables are presented below. Table 4 summarizes the initial characteristics of the test video. Table 5 outlines the criteria for inclusion and exclusion, while Table 6 describes each test video in detail.

Table 4. Test Video Characteristics

Characteristic	**Detail**
Number of videos	100
Length of Videos	5-10 minutes of actual officer-community interaction
Interaction Context	Routine officers-community interactions include both positive and negative interactions.
Included types of police-community contacts	Pedestrian stops, community-initiated calls for service, officer-initiated calls for service

Excluded types of police-community contacts	Traffic stops
Event Exclusion criteria	Videos involving use of force, on-video violence, substantial injury(ies), people under 18, reports of sexual abuse or child abuse Conditions: excessive number of officers or community members, poor lighting or weather, excessively noisy environment

Table 5. Inclusion/Exclusion Criteria

Inclusion Criteria	**Exclusion Criteria**
Daylight and Outdoor conditions	Heavy accents or recordings in languages other than English.
Torso and the faces of community members are visible for as much of the video as possible	Recordings involving violence or use of force resulting in injury
High Resolution	Individuals under 18
Demographic Diversity (race, ethnicity, gender, etc.)	Reports of sexual abuse or child abuse
Minimal noise distortion from vehicle traffic or severe weather	Recordings are part of ongoing investigations
Minimal presence of other people	Excessive number of officers or community members
Close interaction between officer and community member (ideally no less than 10 feet apart when talking)	Traffic Stops
Clear voices (both officer and community member)	Poor lighting or weather
	Noisy environment

Table 6. Description of Videos

VIDEO NUMBER	DESCRIPTION
1	An apartment complex manager reports vagrancy involving a suspicious person.
2	Primary officer contacts a person who reports their license plates were stolen.
3	Officers respond to a family dispute.
4	Officers respond to a landlord-tenant dispute.
5	Officer responds to a domestic disturbance.
6	Officers respond to an intoxicated person lying on the sidewalk.
7	Officers respond to a mental health crisis.
8	Officer responds to a hit-and-run collision.
9	Officers respond to a report of a suspicious vehicle.
10	Primary officer responds to a domestic dispute and related property damage at an apartment complex.

RESULTS AND FINDINGS

A key objective of the project was to compare the evaluations of TrustStat to human coders. As machine learning and artificial intelligence are used to automate tasks that traditionally rely solely on human judgment, the question of human equivalency emerges. That is, how do the judgments of an AI software program, TrustStat in this case, compare to those of humans? The distinction between machine-and human-produced evaluations may be more pronounced when analyzing complex human interactions compared to simpler tasks such as classifying objects. The key concern here is whether AI software can make judgments on procedural justice that are significantly similar to those of coders. Recall the research questions for the project:

1. Are there differences in perceptions of procedural justice between (a) community members, (b) university faculty members and graduate students, and (c) police supervisors?

2. Do procedural justice scores generated by automated video analytics align with scores generated by (a) community members, (b) university faculty members, and (c) police supervisors?

3. Can data from manually coded interactions be used to refine the automated coding algorithms and scoring/weighting procedures?

The first two questions can be answered analytically using statistical tests. The following is a discussion of those tests and the results.

The coding results were downloaded from Qualtrics and cleaned using the R programming language. The survey contained 35 items for each video that captured the dimensions of procedural justice and contextual information about the video, such as the number of community members shown. Additional items were included at the end of the survey to collect group and demographic information from the respondents.

The result was a data frame with 374 variables. The responses for each procedural justice dimension were collapsed together to create one score for each dimension per video and coder. A total procedural justice score for each video and coder were also created. These scores were then aggregated to the group level for analysis. The data included some missing variables when coders needed to provide an answer to every item within the instrument. The average number of missing values across all procedural justice items was small at 0.42. The missing values were imputed with the median score for the video.

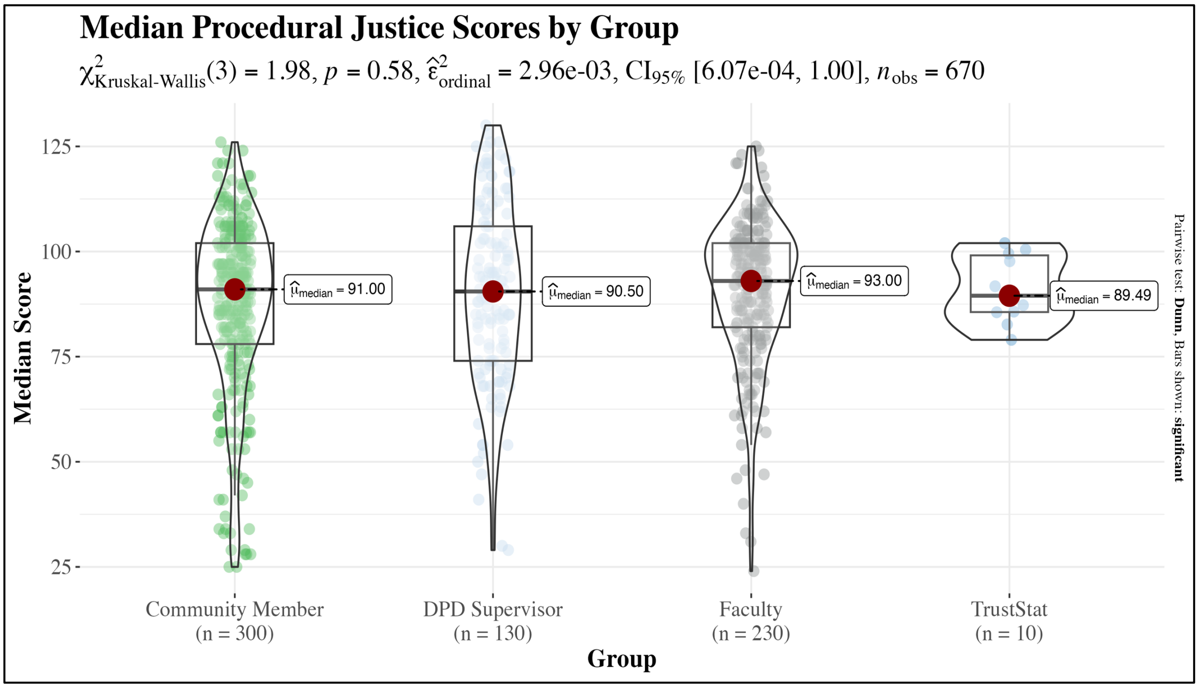
The TrustStat results contained overall ratings for ten videos. The ratings from TrustStat were scaled differently from the coder results, with the results ranging from -0.02 to 0.26. The scores were rescaled based on the upper and lower limits of the coder scores using the rescale function in R (*R: Rescale Variables to a New Range*, n.d.). The data cleaning and processing resulted in a score for each video and coder. Those values served as the basis for the analyses. Comparing group means is typically done using Analysis of Variance or ANOVA. Using ANOVA, though, requires meeting several assumptions about the data. Some assumptions include the data originating from a normal distribution and the distributions having the same variance (*10.2.1 - ANOVA Assumptions | STAT 500*, n.d.). Due to the nature of the sampling, it was unlikely these assumptions would hold for the coding data. To test for normality, a Shapiro-Wilk test was performed, and the p-value was less than 0.05, indicating the normality assumption was violated. The test results were visually confirmed with a Q-Q plot (not shown). The non-parametric alternative to ANOVA is the Kruskal-Wallis test. The omnibus Kruskal-Wallis test focuses on the relative position of the groups or the median position of the groups rather than comparing means as the ANOVA does (Hollander et al., 2014). For this project, the concern is not the specific values of the procedural justice evaluations of coders and groups but how close they are to one another. Knowing the relative position of groups of coders can be used to test the research hypotheses and understand if differences exist between the groups of coders.

The first research question of whether the groups of coders (community members, university faculty members, and police supervisors) differ in their respective ratings can be stated as research hypothesis one.

H1: The groups of human coders will evaluate procedural justice differently.

Given the varying perceptions of police among different groups in society discussed in the literature review and the differences among coder groups characteristics, it was expected that the coders comprised of community members, faculty members, and police supervisors would differ when evaluating the use of procedural justice. For example, police supervisors may be more lenient when evaluating officers. Conversely, they may have higher standards when evaluating due to their additional experience. The results from the Kruskal-Wallis test shown in Figure 1 do not support this hypothesis. The median scores from all three coder groups range from 90.5 to 93.0, with differences not reaching statistical significance using the Kruskal-Wallis test. The other feature to note from Figure 1 is the shape of the plots for each group. Overall, the faculty members had the least variation in the scores. Contrast that with the community members, with a wider range of scores.

Figure 1. Group Medians and Omnibus Test



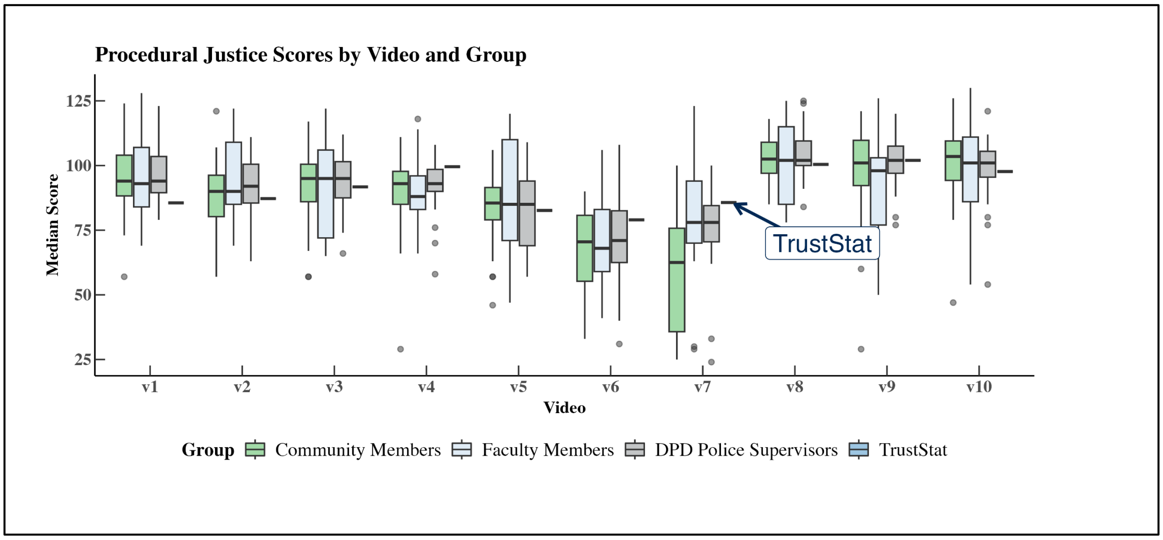
The second research question is whether TrustStat can evaluate BWC footage similarly to human coders. The formal hypothesis for this research question is as follows:

H2: TrustStat evaluations of procedural justice will be consistent with the evaluations of coders.

The rationale for research hypothesis two is that although human interactions are complex and context-dependent, conversations contain objective data that can be used to determine a number of important things. Of course, the team is concerned with officers’ use of procedural justice for this project. The Polis teams’ expertise in understanding human interactions and constructing computer models allowed them to develop TrustStat into a system that could sense and interpret specific verbal cues that translate into procedural justice. Referring to Figure 1, the median rating for TrustStat was 89.49, similar to the median scores for the other groups. The Kruskal-Wallis did not find a statistically significant difference for the medians or relative positions of the groups, including TrustStat. As an omnibus test, the Kruskal-Wallis test only measures differences among all the groups. Tests for pairwise differences in groups were carried out using a Dunn’s test in R done in conjunction with the Kruskal-Wallis test with no significant findings (Dunn, 1961; Patil, 2021). Hypothesis two is supported as there appears to be no statistically significant difference between coder groups and TrustStat.

Looking at the distribution of scores across coders and groups provides another perspective on evaluating the BWC footage. Figure 2 shows box plots of the median scores for each group and video. The plot shows the same pattern of consistent scoring found in Figure 1, but the boxplot shows that the dispersion of scores was much higher in some videos. In particular, video seven demonstrates a wide range of community members' scores. Other videos, such as four, have a much smaller range of scores.

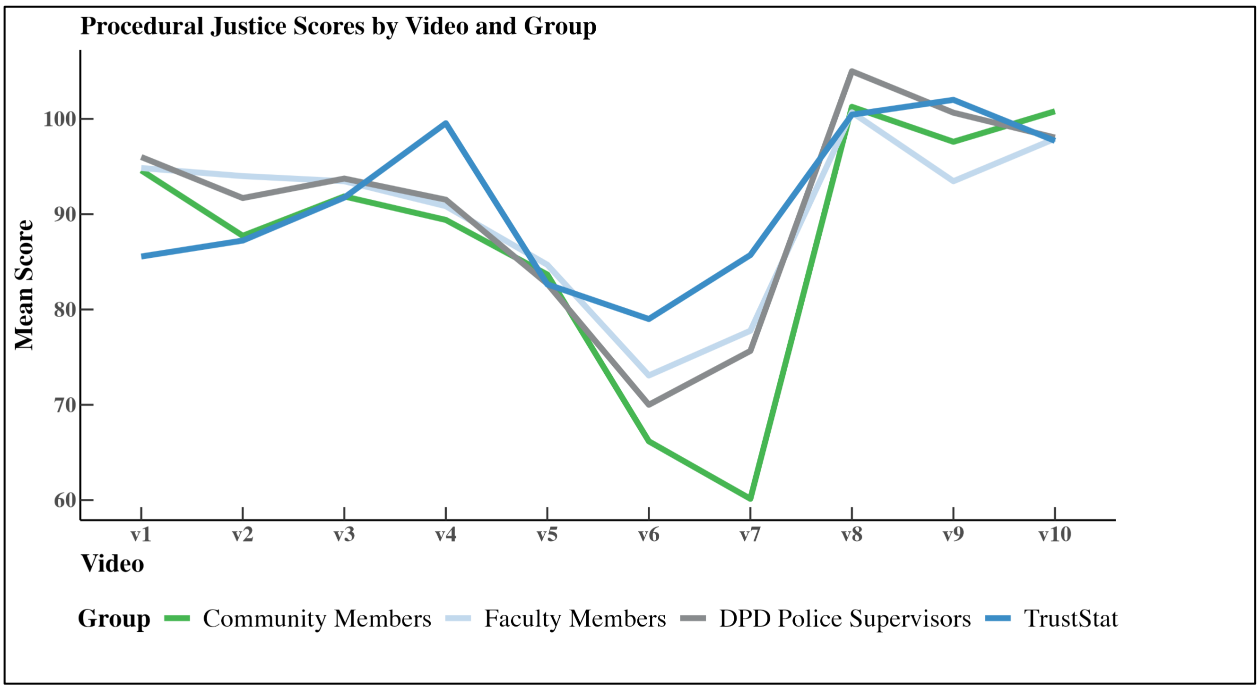
Figure 2. Boxplot of Median Procedural Justice Scores



The final research question was whether the BWC evaluations done by coders could be used to refine and improve the automated evaluations done by TrustStat. In short, the answer is yes. Although this research found that the evaluations were generally consistent, it also demonstrated areas where humans and the algorithm differed. The differences between humans and machines represent an opportunity to raise awareness of those differences and understand the shortcomings of both ways of evaluating the performance of police officers. Those areas will now be discussed.

Figure 3 is a line chart showing the coder group's mean evaluation scores for each video. The plot shows consistency among the groups for most videos. Videos six and seven show the most significant disparity in the mean scores. Some of the most prominent differences came from the dimension of participation, where TrustStat scored videos six and seven much higher than the human coders. For example, TrustStat evaluated video six with a score of 17.2, while the coder groups gave mean scores of 11.3 (community members), 11.5 (police supervisors), and 11.5 (faculty members). Because all coders rated the participation similarly and TrustStat gave a relatively higher score, the software may be more sensitive to cues of involvement than a human watching the interaction. Another explanation may be that humans perceive participation at some level but do not acknowledge it as meaningful because of contextual factors such as tone or inflection.

Figure 3. Mean Procedural Justice Scores by Video and Group



Videos six and seven showed community members in distress because of what appeared to be a mental health issue from the information provided in the video. In both videos, the community member experiencing distress offered little or no communication to the officer(s) on the scene. Moreover, the engagements between the officers and community members in video seven exhibited a heightened sense of tension and anger in the case of one community member. These two videos were likely subjectively different from the other videos in terms of the amount of distress from the community members. It could be that the human coders responded to the emotions differently than TrustStat.

DISCUSSION OF FINDINGS AND IMPLICATIONS

Developing software to evaluate body-worn camera (BWC) footage systematically has the potential to significantly enhance the analysis of police data. Comparing TrustStat to coders yielded several noteworthy discoveries. One key finding is that despite the diverse demographic backgrounds among the coders involved in this study, they rated procedural justice similarly across all ten videos but demonstrated variability within videos. This consistency across videos underscores the reliability of the instrument and its accompanying instructions in producing dependable results. However, for some videos, there were differences within the groups. Examples of this are the variation in faculty members for video five and community members in video seven. More research needs to be done to understand the causes of the variation within groups. Although it is unknown whether groups of coders from different localities will have the same judgments, it does provide optimism that the underlying concepts of procedural justice can be systematically quantified by diverse groups of coders with relatively little training and understanding of police practices (Grahama et al. 2024; Matsumoto et al., 2024).

The second key finding is that TrustStat evaluated the videos similarly to average coders. Despite some differences in the nuances of procedural justice, the software produced equivalent judgments of procedural justice. This finding has vast implications for law enforcement. If procedurally just encounters minimize risk to officers through voluntary compliance and foster police legitimacy, systematically quantifying the levels of procedural justice can transform the way officer performance is managed. Rather than relying on subjective performance assessments from supervisors or using indicators such as complaints from citizens, officers can be assessed on observable criteria that relate to how they directly interact with the community. Enhancing the relationship between officers and the community has long been a priority for communities, scholars, and law enforcement agencies.

Software such as TrustStat has the potential to be incorporated into early warning systems to flag potentially aberrant and troublesome behaviors captured by BWC. More ambitiously, however, AI analysis has the potential to identify at scale the patterns of behavior and language that correlate with the positive, peaceful outcomes that are the norm in most police-community interactions. Currently, all footage lies fallow stored on servers, only to be viewed if a complaint or controversy arises. Using software in a proactive manner could transform the current paradigm of officer performance. Rather than anecdotal accounts of good or bad officer performance, organizations could look more holistically at patterns of behaviors throughout a performance period and anchor the evaluations on interactions with the community. Automated, large-scale BWC analysis provides an opportunity to truly judge officers on how they interact with the community in ways that are anchored in data rather than on low-frequency events such as complaints or commendations from the community.

LIMITATIONS

This research supports a systematic approach to analyzing BWC footage. It provides a foundation for using AI software to evaluate BWC footage but has inherent limitations. The primary limitations were the exclusion rules for videos, the convenience sample for coders, and the lack of validation from the lived experiences of community members. These limitations will be discussed in the following sections.

As mentioned in previous sections, selecting the final included test videos was limited, as predetermined by IRB guidelines, to protect human subjects and reduce harm. Videos that included protected subjects displayed use of force or were unclear visually and in audio were excluded from further selection. Videos that involved more than one responding officer or multiple community members on scene were excluded to further focus on the initial officer-community member interaction. From the sample of 100 unredacted video recordings, ten were purposely selected as test videos. After further review, the project team selected ten final videos to be edited as part of the test sample.

Additional limitations included the voluntary participation of Dallas community members, faculty, and police personnel. The project team utilized several recruiting techniques to encourage participation in the study, such as email communication and in-person outreach. The project team routinely encouraged participation through follow-up communication. Due to the extensive time it took to complete the survey, participants were reimbursed via electronic gift cards. While there was equal participation from both community members and faculty, neither group should be considered the "voice" of the community or university faculty. Additionally, there were challenges with recruiting additional police personnel. As many departments across the country experience resource constraints due to adequate staffing levels, this also appeared to be a challenge for the DPD.

Related to the selection of coders is the project's geographical exclusivity. The videos, coders, and larger context are centered in Dallas, Texas. The BWC footage was from the DPD, and the coders were all from the Dallas area. This narrow geographic focus could bias the findings, especially compared to a more regional or national approach. For example, public opinions of the DPD in the Dallas area could be better or worse than sentiments regarding departments in other communities. Recent incidents, good or bad, could also impact the views and, ultimately, evaluations of officers. Moreover, the recruitment of coders was not random and should be considered a convivence sample. Because of the nature of the sample, the coders who volunteered may have different views than a randomly selected community member.

A final and important limitation is the absence of the accounts from the community members' interaction with the officers. Having the coders and software take an omniscient view of the interactions and then make judgments about procedural justice does not account for how the community member feels during the interaction. What may seem like a procedurally just interaction may be objectionable or offensive to the community member experiencing it. The difference in perspective may be amplified by the socio-economic differences between the coders and the community members experiencing the police in their day-to-day lives.

IMPLICATIONS FOR FUTURE RESEARCH

This study found that disparate groups of coders and AI software developed to evaluate the use of procedural justice produce similar evaluations. Two critical areas in this line of research need further investigation. One of those areas is to broaden the scope of the study. The pool of coders should be expanded to incorporate more diverse geographic areas and coders. Ideally, the selection of coders would be randomly drawn from the city or jurisdiction where the BWC footage originates for better generalization of results. A corollary to this broadening is to include a more diverse set of videos with greater complexity to test the respective perceptual abilities of human coders and TrustStat AI software.

Another focus of future research is to incorporate mechanisms to capture the perceptions of the community members interacting with officers in the BWC footage. As mentioned, there is a need to understand how an objective evaluation of an interaction, whether judged by coders or software, correlates to the lived experiences of the community members and, importantly, also the officers. Such research could be done using contact surveys, enabling community members to rate their interactions with officers. Once the interaction is rated by the involved community members, the BWC footage from the interaction could be automatically flagged for review by software or human reviewers software. Similarly, officers could log into TrustStat and directly input their own ratings. Just as this project has compared the judgments of TrustStat to coders, similar comparisons are needed with community members experiencing the interaction.

CONCLUSIONS

The findings from this study underscored the potential of BWC data and advanced AI analytics to transform the evaluation of police interactions with the community. The project team has shown that TrustStat, an automated video analysis tool, can generate procedural justice evaluations comparable to those conducted by human coders. This consistency across diverse groups—community members, university faculty, and police supervisors—highlights the robustness of the procedural justice framework and TrustStat's efficacy.

The implications for law enforcement are significant. By integrating tools like TrustStat, agencies can systematically and objectively assess officer performance based on longitudinal patterns of community interactions rather than relying solely on subjective reports, random audits, or infrequent incidents. This transformational shift in the use of BWC data could help enhance accountability, improve training, and ultimately foster greater trust and police legitimacy within the community.

However, this study also identified key limitations, such as the geographical concentration of data from Dallas, Texas, and the non-random selection of coders. These factors suggest the need for broader, more diverse datasets and randomized coder selection in future research to ensure the generalizability of the findings. Additionally, incorporating direct feedback from community members involved in the interactions remains crucial for a holistic understanding of procedural justice.

The advancements in BWC analytics represented by TrustStat offer a promising avenue for enhancing police-community relations by leveraging technology to provide data-driven insights, law enforcement agencies can better understand and improve their interactions with the public, paving the way for a more just and equitable policing system. Future research should continue to refine these tools and expand their application to ensure that the benefits of procedural justice are realized across diverse communities and contexts.

ARTIFACTS

Artifact 1. Coding Instrument

S054 - Body Worn Camera Coding Instrument

Start of Block: Block 7

Analysis of Body Worn Camera Recordings: Measuring Police Implementation of Procedural Justice SSO DATA COLLECTION INSTRUMENT
Project funded by the National Institute of Justice

Last updated 4/12/2023 9:46:00 AM

Page Break

The National Policing Institute is assessing the use of procedural justice by police officers during interactions with the community. To evaluate those interactions, you will be asked to watch a series of body-worn camera (BWC) videos and then take a survey comprised of 35 items capturing critical aspects of procedural justice.

There are minimal risks associated with this research. Some of the risks include the potential loss of confidentiality from other participants and distress from watching BWC videos. Your participation is voluntary and you may choose not to participate in this research study or withdraw your consent at any time. You will NOT be penalized in any way should you choose not to participate or withdraw. The research team will do everything we can to protect your privacy. As part of this effort, your personal identity will not be collected. If you wish to continue, please click proceed.

○ Proceed (4)

Page Break

INSTRUCTIONS

The National Policing Institute is assessing the use of procedural justice by police officers during interactions with the community. To evaluate those interactions, you will be asked to watch a series of body-worn camera (BWC) videos and then take a survey comprised of 35 items capturing critical aspects of procedural justice.

You will first view an entire BWC video from beginning to end. Then begin the survey while watching the video again, starting and stopping it as needed to respond to items in the survey. The videos may show multiple officers. Your focus will be on the interaction between the BWC-wearing officer, referred to as the "primary officer" in the survey. Since this officer is wearing a body camera, the point of view will be first person, and you will not see the officer. The videos may also have multiple community members in the video. For any survey items related to the community member, your reference is to the community member the officer interacts with most. This community member will be referred to as the "primary community member." The distinction between the primary community member and other community members should be evident based on the direction and volume of the interactions.

When answering the questions in the survey, you should consider the interaction as a whole. While the primary officer may demonstrate different behaviors throughout the scenario, you should think about the totality of the officer's behavior when evaluating procedural justice. For example, an officer may briefly demonstrate disrespect towards a community member, but for the majority of the interaction is respectful.

Please ask a research team member present if you have any questions or concerns when responding to survey items.

End of Block: Block 7

Start of Block: Block 8

Video Number

____

End of Block: Block 8

Start of Block: Scenario Context

Q1 Besides the primary officer, how many officers were visible at the scene?

- ○ None (primary officer only) (1)
- ○ One (2)
- ○ Two (3)
- ○ Three (4)
- ○ Four or more (5)
- ○ Unable to determine (6)

Q2 Besides the primary community member, how many community members were at the scene?

- ○ None (primary community member only) (1)
- ○ One (2)
- ○ Two (3)
- ○ Three (4)
- ○ Four or more (5)
- ○ Unable to determine (6)

Q3 What was the primary community member's gender?

- ○ Male (1)
- ○ Female (2)
- ○ Unable to determine (3)

Q4 What was the primary community member's age?

- Adult (18-44) (1)
- Middle-aged (45-59) (2)
- Senior (60 and above) (3)
- Unable to determine (4)

Q5 What was the primary community member's race?

- White (1)
- Black or African American (2)
- American Indian or Alaska Native (3)
- Asian (4)
- Native Hawaiian or Pacific Islander (5)
- Two or more races (6)
- Unable to determine (7)

Q6 What was the primary community member's ethnicity?

- Hispanic or Latino (1)
- Not Hispanic or Latino (2)
- Other (3)
- Unable to determine (4)

End of Block: Scenario Context

Start of Block: Community Member Information

Q7 The primary community member appeared to be under the influence of alcohol or other drugs.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q8 The primary community member showed signs of a behavioral health condition.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q9 The primary community member <u>showed signs of physical injury or illness</u>.

- ◯ Strongly disagree (1)
- ◯ Disagree (2)
- ◯ Somewhat disagree (3)
- ◯ Neither agree nor disagree (4)
- ◯ Somewhat agree (5)
- ◯ Agree (6)
- ◯ Strongly agree (7)

Q10 The primary community member showed signs of extreme emotions (e.g., sobbing, screaming).

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q11 The primary community member was in conflict with another community member on the scene.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Community Member Information

Start of Block: Neutrality

Q12 The primary officer indicated they would seek all viewpoints about the matter.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q13 The primary officer indicated that they would not decide what to do until they had gathered all the necessary information.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q14 The primary officer indicated that his/her decision in this situation was not influenced by the personal characteristics (race, age, sex) of anyone present.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Somewhat disagree (3)
- ○ Neither agree nor disagree (4)
- ○ Somewhat agree (5)
- ○ Agree (6)
- ○ Strongly agree (7)

Q15 The primary officer explained why the police became involved in the situation.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Somewhat disagree (3)
- ○ Neither agree nor disagree (4)
- ○ Somewhat agree (5)
- ○ Agree (6)
- ○ Strongly agree (7)

Q16 The primary officer explained to the community member why they chose to resolve the situation as they did.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Neutrality

Start of Block: Participation

Q17 The primary officer asked or told the primary community member to provide their information or viewpoint.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Somewhat disagree (3)
- ○ Neither agree nor disagree (4)
- ○ Somewhat agree (5)
- ○ Agree (6)
- ○ Strongly agree (7)

Q18 The primary officer allowed the community member to provide information or viewpoint without interrupting.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q19 The primary officer <u>expressed interest in the information or viewpoint</u>.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Participation

Start of Block: Trust

Q20 The primary officer asked the community member about their well-being or asked other individuals at the scene in a way that the community member observed.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q21 The primary officer offered comfort or reassurance to the community member.

- ◯ Strongly disagree (1)
- ◯ Disagree (2)
- ◯ Somewhat disagree (3)
- ◯ Neither agree nor disagree (4)
- ◯ Somewhat agree (5)
- ◯ Agree (6)
- ◯ Strongly agree (7)

Q22 The primary officer provided or promised to exert control or influence over another person for the community member.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q23 The primary officer indicated a report would be filed for the community member.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Somewhat disagree (3)
- ○ Neither agree nor disagree (4)
- ○ Somewhat agree (5)
- ○ Agree (6)
- ○ Strongly agree (7)

Q24 The primary officer acted or promised to act on behalf of the community member with a government agency or other organization.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q25 The primary officer provided/arranged or promised to provide/arrange physical assistance to the community member.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q26 The primary officer provided or promised to provide advice on how the community member could handle the situation or deal with the problem.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Trust

Start of Block: Respect

Q27 The primary officer used a greeting or introduction when speaking with the community member.

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q28 The primary officer <u>used respectful or formal titles when addressing the community member (e.g., Sir, Ma'am, Mr., Ms.)</u>

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Q29 The primary officer <u>used swear words or disrespectful language when speaking with the community member (e.g., "You got a shit ton of warrants".)</u>

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Neither agree nor disagree (4)

○ Somewhat agree (5)

○ Agree (6)

○ Strongly agree (7)

Q30 When considering the entire interaction between the primary officers and the primary community member, the primary officer was respectful.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Neither agree nor disagree (4)

○ Somewhat agree (5)

○ Agree (6)

○ Strongly agree (7)

End of Block: Respect

Start of Block: Coder Information

Q31 Group affiliation

○ Community Member (1)

○ Police Employee (2)

○ University Employee (3)

Q32 What is your gender?

- Male (1)
- Female (2)
- Transgender (3)
- Non-binary (4)
- Do not wish to disclose (5)

Q33 What was your age (in years) on your last birthday?

____

Q34 What is your ethnic origin or race?

- ○ Non-Hispanic White (1)
- ○ Hispanic/Latino (2)
- ○ Black/African American (3)
- ○ Native American/Indian (4)
- ○ Asian (5)
- ○ Pacific Islander (6)
- ○ Other (please specify) (7) ____________________

Q35 What is your level of education?

- ○ Less than high school (1)
- ○ High school graduate or the equivalent (e.g., GED) (2)
- ○ Some college (no degree) (3)
- ○ 2 year degree (4)
- ○ 4 year degree (5)
- ○ Master's degree or higher (6)

End of Block: Coder Information

Artifact 2. Recruitment Scripts

Dallas PD Supervisor Recruitment Script

Hello, my name is___________________. I am a faculty member and researcher with the University of North Texas at Dallas in the Department of Criminal Justice and Sociology. The reason for my email is to seek your assistance in a study that I and my colleagues are conducting that examines perceptions of procedural justice. After watching Officer Body Worn Camera videos, you will be asked to answer questions about procedural justice. The video and completion of the accompanying surveys should take a total of three hours to complete. Your participation is fully voluntary, and your participation or refusal to participate will not in any way affect your standing in the department. We can provide any additional information regarding this study you deem necessary or needed to decide on our request.

Although minimal, there is a certain amount of risk that is involved with participating in this research study. If you feel uncomfortable about any of these risks, discuss your concerns with the research team. Our contact information is included in this sheet. We understand that it is not possible to identify all potential risks in survey research, but we believe that reasonable safeguards have been taken to minimize both known and unknown potential risks associated with this study. This study involves informational risk, such as loss of confidentiality, meaning that information we collect about you could be accessed by someone not authorized to see it. However, we will work hard to protect the information we collect about you and to keep it private. Your name will not ever be recorded on the study answer forms. You do not have to enroll in this study and may withdraw your consent to participate at any time before completion. Since no identifiers will be collected as part of the study procedures, there will be no way for the research team to retrieve individual responses once they are submitted electronically.

Community Member Recruitment Script

Hello, my name is___________________. I am a faculty member and researcher with the University of North Texas at Dallas in the Department of Criminal Justice and Sociology. The reason for my email is to seek your assistance in a study that I and my colleagues are conducting that examines perceptions of procedural justice. After watching Officer Body Worn Camera videos, you will be asked to answer questions about procedural justice. The video and completion of the accompanying surveys should take a total of three hours to complete. Your participation is fully voluntary, and your participation or refusal to participate will not in any way affect you. We can provide any additional information regarding this study you deem necessary or needed to decide on our request.

Although minimal, a certain amount of risk is involved with participating in this research study. If you feel uncomfortable about any of these risks, discuss your concerns with the research team. Our contact information is included in this sheet. We understand that it is not possible to identify all potential risks in survey research, but we believe that reasonable safeguards have been taken to minimize both known and unknown potential risks associated with this study. This study involves informational risk, such as loss of confidentiality, meaning that information we collect about you could be accessed by someone not authorized to see it. However, we will work hard to protect the information we collect about you and to keep it private. Your name will not ever be recorded on the study answer forms. You do not have to enroll in this study and may withdraw your consent to participate at any time before completion. Since no identifiers will be collected as part of the study procedures, there will be no way for the research team to retrieve individual responses once they are submitted electronically.

Artifact 3. Datasets Created

The dataset generated in this project comprises the coding outcomes from the sessions conducted at the University of North Texas Dallas. At the completion of the project, the data will be archived as agreed upon in the project award.

Artifact 4. Dissemination Activities

This study's dissemination includes publication in academic journal articles, trade magazines such as *Police Chief*, industry websites for law enforcement, such as *Police1.com*, and research events and conferences for law enforcement practitioners and academics.

Artifact 5. Annotated Dimensions.

	Rating Structure Used	
Annotated Dimension	Was this evident in the encounter? *(Yes, No)*	Frequency/Duration *(Never, Very little, Some of the time, Most of the time, All of the time)*
<u>Participation</u>		
PO asked PCM questions.	x	
PCM asked PO questions.	x	
PO addressed PCM's questions; PO provided information/viewpoint/relevant facts for PCM.	x	x
PCM addressed PO's questions; PCM provided information/viewpoint/relevant facts for PO.	x	x

PO allowed PCM to speak/provide info without interrupting.		x
PCM allowed PO to speak/provide info without interrupting.		x
PO expressed interest/desire in PCM's information/viewpoint.	x	

Neutrality

PO spoke in a way that was easy to understand.		x
PCM spoke in a way that was easy to understand.		x
PO explained goal of interaction/why PO chose to resolve the situation as they did.	x	
PO made sure they understood what the PCM was communicating.	x	
PO indicated they would seek all viewpoints about the matter.	x	
PO communicated with PCM while carrying out a task.	x	

Respect

PO behaved/spoke respectfully to PCM during encounter.		x
PCM behaved/spoke respectfully to PO during encounter.		x
PO behaved/spoke disrespectfully to PCM during encounter.		x
PCM behaved/spoke disrespectfully to PO during encounter.		x
PO stood close enough for communication, but with enough distance to ensure safety and to respect personal boundaries.		x

PCM stood close enough for communication, but with enough distance to ensure safety and to respect personal boundaries.		x
PO used profanity in light/casual/friendly way when speaking to PCM.		x
PCM used profanity in light/casual/friendly way when speaking to PO.		x
PO used profanity in hostile/demeaning/unfriendly way when speaking to PCM.		x
PCM used profanity in hostile/demeaning/unfriendly way when speaking to PO.		x
PO was adaptable (changed their behavior to adjust to changing conditions, switched tactics when original tactics didn't work, recognized and repaired errors or misunderstandings).		x
PCM was adaptable (changed their behavior to adjust to changing conditions, switched tactics when original tactics didn't work, recognized and repaired errors or misunderstandings).		x
PO was hostile/aggressive to PCM.		x
PCM was hostile/aggressive to PO.		x
PO was upset/agitated with PCM.		x
PCM was upset/agitated with PO.		x
PO physically attacked/assaulted PCM/someone else in PCM's presence.	x	
PCM physically attacked/assaulted PO/someone else in PO's presence.	x	
PO was polite to PCM.		x
PCM was polite to PO.		x
PCM was compliant/followed orders/did not resist.		x

Trust

PO offered advice/assistance/resources for PCM.	x	
PCM offered advice/assistance/resources for PO.	x	
PO behaved/spoke in a casual/friendly way that helped build rapport and reduce tension.		x
PCM behaved/spoke in a casual/friendly way that helped build rapport and reduce tension.		x
PO behaved/spoke in a calm manner that helped reduce tension.		x
PCM behaved/spoke in a calm manner that helped reduce tension.		x
PO demonstrated concern for PCM's safety and welfare.	x	x
PO attempted to establish common ground with PCM.	x	
PO showed empathy/humanity for PCM.	x	x
PO acknowledged/recognized PCM's viewpoints (e.g., I understand your point; Makes sense; I see, etc).	x	
PO expressed regretful acknowledgment(s) to PCM (e.g., sorry, oops, excuse me, etc).	x	
PO expressed gratitude/appreciation to PCM (e.g., thank you, i appreciate it, etc.).	x	
PO requested action for self from PCM (let me take a step back for a minute, can i take you over here? etc).	x	
PCM felt victimized by PO.	x	
Rate PO's tone	Very negative tone, Negative tone, Somewhat negative tone, Neutral tone, Somewhat positive tone, Positive tone, Very positive tone	

Rate PCM's tone	Very negative tone, Negative tone, Somewhat negative tone, Neutral tone, Somewhat positive tone, Positive tone, Very positive tone
Rate PO's overall performance on procedural justice.	Below average, Average, Excellent

Note. PO = primary police officer; PCM = primary community member

Artifact 6. References

10.2.1—ANOVA Assumptions | STAT 500. (n.d.). Retrieved June 24, 2024, from https://online.stat.psu.edu/stat500/lesson/10/10.2/10.2.1

Alpert, G., & Dunham, R. (2004). Understanding Police Use of Force: Officers, Suspects, and Reciprocity. Understanding Police Use of Force: Officers, Suspects, and Reciprocity. 1-191. 10.1017/CBO9780511499449.

Ariel, B., Farrar, W. A., & Sutherland, A. (2014). The effect of police body-worn cameras on use of force and citizens' complaints against the police: A randomized controlled trial. *Journal of Quantitative Criminology*, *31*(3), 509-535. https://doi.org/10.1007/s10940-014-9236-3

Barboza, G. E. (2012). Group consciousness, identity and perceptions of unfair police treatment among Mexican Americans. *Policing: An International Journal of Police Strategies & Management*, *35*(3), 505-527.

Benedict, M. E., Brown, W. A., & Bower, B. T. (2000). Perceptions of police legitimacy in the United States: A test of competing theoretical perspectives*. Criminology, 38(3), 647-678.

Boivin, R., Faubert, C., Gendron, A., & Poulin, B. (2020). Explaining the body-worn camera perspective bias. *Journal of Qualitative Criminal Justice & Criminology, 9*(1), 1-28.

Brown, W. A., & Benedict, M. E. (2002). Community context, racial/ethnic composition, and perceptions of police legitimacy: A multi-level analysis. Police Quarterly, 5(1), 1-24.

Callanan, V. J., & Rosenberger, J. S. (2011). Media and public perceptions of the police: Examining the impact of race and personal experience. *Policing & Society*, *21*(2), 167-189.

Copeland, C., del Carmen, A., & Semukhina, O. B. (2022). Revisiting the Ferguson effect: Law enforcement perception of recruitment in the post George Floyd era. *International Journal of Police Science & Management*, *24*(3), 261-272.

Coppola, M. (2010). TECHBeat - NIJ National Law Enforcement and Corrections Technology Center (NLECTC). *JustNet*. https://www.ojp.gov/pdffiles1/nij/nlectc/233930.pdf

Crow, M. S., Snyder, J. A., Crichlow, V. J., & Smykla, J. O. (2017). *Community Perceptions of police body-worn cameras. Criminal Justice and Behavior 44(4), 5889-610.*

Dignan, L. (2019). Axon moves 20 PB of data, Evidence.com to Microsoft Azure. *ZDNET*. https://www.zdnet.com/article/axon-moves-20-pb-of-data-evidence-com-to-microsoft-azure/

Dunn, O. J. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association*, *56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Farooq, U. (2024). Police turn to AI to review bodycam footage. *ProPublica*. https://www.propublica.org/article/police-body-cameras-video-ai-law-enforcement

Feeney, M. (2015). *Watching the Watchmen: Best Practices for Police Body Cameras. Cato Institute. Washington, D.C*. Shibboleth Authentication Request. https://www-cato-org.proxy.cc.uic.edu/publications/policy-analysis/watching-watchmen-best-practices-police-body-cameras

Forgas, J. P., & Bond, M. H. (1985). Cultural influences on the perception of interaction episodes. *Personality and social psychology bulletin*, 11(1), 75-88.

Fourcade, M., Lande, B., & Schofer, E. (2016). Political space and the space of polities: Doing politics across nations. *Poetics*, 55, 1-18;

Friis, C. B., Liebst, L. S., Philpot, R., & Lindegaard, M. R. (2020). Ticket inspectors in action: Body-worn camera analysis of aggressive and nonaggressive passenger encounters. *Psychology of Violence*, *10*(5), 483–492. https://doi.org/10.1037/vio0000276

Gonzales, A. R., & Cochran, D. Q. (2017). *Police-worn body cameras: An antidote to the Ferguson effect* (82 (2): 299-338). Missouri Law Review.

Goodison, S. E., Brooks, C., & Bureau of Justice Statistics. (2023). Local Police Departments, Procedures, Policies, and Technology, 2020 – Statistical tables. In *Bureau of Justice Statistics*. https://bjs.ojp.gov/document/lpdppt20st.pdf

Grahama, B. A., Brown, L., Chochlakis, G., Dehghani, M., Delerme, R., Friedman, B., Narayan, S. (2024). A Multi-Perspective Machine Learning Approach to Evaluate Police-Driver Interaction in Los Angeles. arXiv preprint arXiv:2402.01703.

Griffith, C. E., & Foley, A. J. (2020). The relationship between global perceptions of police in the United States and citizen satisfaction with local police. *Sociation, 19*(2), 37-47.

Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Nonparametric statistical methods* (Third edition). John Wiley & Sons, Inc.

Huggins, C. M. (2012). Traffic stop encounters: Officer and Community member, individual, or public race and perceptions of police propriety. *American Journal of Criminal Justice*, *37*, 92-110.

Jameel, L., & Bunn, S. (2015). Body-Worn Video in UK Policing. *Parliamentary Office of Science and Technology, London*. https://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PB-0014

James, V. J., & Lee, D. R. (2015). Through the looking glass: Exploring how college students' perceptions of the police influence sexual assault victimization reporting. *Journal of interpersonal violence*, *30*(14), 2447-2469.

Jennings, W. G., Fridell, L. A., & Lynch, M. D. (2014). Cops and cameras: Officer perceptions of the use of body-worn cameras in law enforcement. *Journal of Criminal Justice*, *42*(6), 549-556. https://doi.org/10.1016/j.jcrimjus.2014.09.008

Jennings, W. G., Lynch, M. D., & Fridell, L. A. (2015). Evaluating the impact of police officer body-worn cameras (BWCs) on response-to-resistance and serious external complaints: Evidence from the Orlando police department (OPD) experience utilizing a randomized controlled experiment. *Journal of Criminal Justice*, *43*(6), 480-486. https://doi.org/10.1016/j.jcrimjus.2015.10.003

Jonathan-Zamir, T., Mastrofski, S. D., & Moyal, S. (2015). Measuring Procedural Justice in Police-Citizen Encounters. *Justice Quarterly*, 32(5), 845–871. https://doi.org/10.1080/07418825.2013.845677

Koslicki, W. M., Makin, D. A., & Willits, D. (2019). When no one is watching: evaluating the impact of body-worn cameras on use of force incidents. *Policing and society*. *30*(5), 569–582. https://doi.org/10.1080/10439463.2019.1576672

Lai, Y. L., & Zhao, J. S. (2010). The impact of race/ethnicity, neighborhood context, and police/citizen interaction on residents' attitudes toward the police. *Journal of Criminal Justice*, *38*(4), 685-692.

Livingston, J. D., Desmarais, S. L., Verdun-Jones, S., Parent, R., Michalak, E., & Brink, J. (2014). Perceptions and experiences of people with mental illness regarding their interactions with police. *International journal of law and psychiatry*, *37*(4), 334-340.

Lum, C., Stoltz, M., Koper, C. S., & Scherer, J. A. (2019). Research on body-worn cameras. *Criminology & Public Policy*, *18*(1), 93–118. https://doi.org/10.1111/1745-9133.12412

Makin, D. A. (2017). Avoiding the technological panacea: The case of the body worn camera. *Criminal justice technology in the 21st century*, 86-102.

Makin, D. A., Willits, D. W., & Brooks, R. (2019). Systematic social event modeling: a methodology for analyzing body-worn camera footage. *International Journal of Social Research Methodology*, *24*(2), 163-176.

Mastrofski, S. D., Jonathan-Zamir, T., Moyal, S., & Willis, J. J. (2015). Predicting Procedural Justice in Police–Citizen Encounters. *Criminal Justice and Behavior*, *43*(1), 119–139. https://doi.org/10.1177/0093854815613540

Matsumoto, D., Wilson, M., Helmy, M., García, J. M. G., & Vasquez, A. N. A. (2024). Cultural Similarities and Differences in Judgments of Rapport. *Journal of Cross-Cultural Psychology*, 55(2), 127-144.

Mazerolle, L., Antrobus, E., Bennett, S., & Tyler, T. R. (2013). Shaping Citizen Perceptions of Police Legitimacy: A Randomized Field Trial of Procedural Justice: Shaping Citizen Perceptions of Police. *Criminology*, *51*(1), 33–63. https://doi.org/10.1111/j.1745-9125.2012.00289.x

McCluskey, J. D., & Reisig, M. (2017). Explaining procedural justice during police-suspect encounters. *Policing*: *An International Journal of Police Strategies and Management*, *40*(3), 574–586. https://doi.org/10.1108/pijpsm-06-2016-0087

McCluskey, J., Uchida, C. D., Feys, Y., & Solomon, S. E. (2023). Developing and Conducting an SSO Study: A How-to-Guide for Social Science Research. In *SpringerBriefs in criminology* (pp. 17–46). https://doi.org/10.1007/978-3-031-31482-7_3

McCluskey, J. D., Uchida, C. D., Solomon, S. E., Wooditch, A., Connor, C., & Revier, L. (2019). Assessing the effects of body-worn cameras on procedural justice in the Los Angeles Police Department. *Criminology*, *57*(2), 208–236. https://doi.org/10.1111/1745-9125.12201

Morrow, W. J., Vickovic, S. G., & Shjarback, J. A. (2021). Motivation to enter the police profession in the post-Ferguson era: An exploratory analysis of procedural Justice. *Criminal Justice Studies*, *34*(2), 135-155.

Nadal, K. L., Davidoff, K. C., Allicock, N., Serpe, C. R., & Erazo, T. (2017). Perceptions of police, racial profiling, and psychological outcomes: A mixed methodological study. *Journal of Social Issues*, *73*(4), 808-830.

National Institute of Justice, (2012). Primer on Body-Worn Cameras for law enforcement. (No.242268). *U.S. Department of Justice, National Institute of Justice*. https://nij.ojp.gov/library/publications/primer-body-worn-cameras-law-enforcement

National Law Enforcement and Corrections and Technology Center. (2010). *Officer Worn Cameras: Expand Point of View*. *Fall*. https://www.ojp.gov/pdffiles1/nij/nlectc/233930.pdf

Patil, I. (2021). Visualizations with statistical details: The "ggstatsplot" approach. *Journal of Open Source Software*, *6*(61), 3167.

Payne, B. K., & Gainey, R. R. (2007). Attitudes about the police and neighborhood safety in disadvantaged neighborhoods: The influence of criminal victimization and perceptions of a drug problem. *Criminal Justice Review*, *32*(2), 142-155.

Peck, J. H. (2015). Minority perceptions of the police: A state-of-the-art review. *Policing: an international journal of police strategies & management*, *38*(1), 173-203.

Goetschel, M., & Peha, J. M. (2017). Police perceptions of body-worn cameras. *American Journal of Criminal Justice, 42*, 698-726.

Peyton, K., Sierra-Arévalo, M., & Rand, D. G. (2019). A field experiment on community policing and police legitimacy. *Proceedings of the National Academy of Sciences*, *116*(40), 19894-19898.

Piza, E. L., & Sytsma, V. A. (2022). The impact of suspect resistance, informational justice, and interpersonal Justice on time until police use of Physical Force: a survival analysis. *Crime & Delinquency 70*(1), 3–28. https://doi.org/10.1177/00111287221106947

Piza, E. and Sytsma, V. (2023). Video Data Analysis of Body-Worn Camera Footage: A Practical Methodology in Support of Police Reform. In Craig, M. and Blount-Hill, K. (Eds.) Justice and Legitimacy in Policing: Transforming the Institution (pp. 59-75). Abingdon, Oxon, UK: Routledge Press.

Prabhakaran, V., Griffiths, C., Su, H., Verma, P., Morgan, N., Eberhardt, J. L., & Jurafsky, D. (2018). Detecting institutional dialog acts in police traffic stops. *Transactions of the Association for Computational Linguistics*, *6*, 467–481. https://doi.org/10.1162/tacl_a_00031

President's Task Force on 21st Century Policing (2015). *Final Report of the President's Task Force on 21st Century Policing. Office of Community Oriented Policing Services, Washington, DC.*

Reisig, M. D., & Parks, R. B. (2003). Neighborhood context, police behavior and satisfaction with police. *Justice Research and Policy*, *5*(1), 37-65.

Rengifo, A. F., & Fratello, J. (2015). Perceptions of the police by immigrant youth: Looking at stop-and-frisk and beyond using a New York City sample. *Youth Violence and Juvenile Justice*, *13*(4), 409-427.

Schuck, A. M., & Rosenbaum, D. P. (2005). Global and neighborhood attitudes toward the police: Differentiation by race, ethnicity and type of contact. *Journal of Quantitative Criminology*, *21*, 391-418.

Schuck, A. M., Rosenbaum, D. P., & Hawkins, D. F. (2008). The influence of race/ethnicity, social class, and neighborhood context on residents' attitudes toward the police. *Police Quarterly*, *11*(4), 496-519.

Skaggs, S. L., Harris, C., & Montgomery, L. (2022). The impact of police–community relations: Recruitment and retention concerns of local police agencies. *Policing: A Journal of Policy and Practice*. 16(3)462-475.

Skogan, W. G. (2006). Asymmetry in the impact of encounters with police. *Policing & Society*, *16*(02), 99-126.

Smykla, J. O., Crow, M. S., Crichlow, V. J., & Snyder, J. A. (2015). Police Body-Worn Cameras: Perceptions of Law Enforcement Leadership. *American Journal of Criminal Justice*, *41*(3), 424–443. https://doi.org/10.1007/s12103-015-9316-4

Sprott, J. B., & Doob, A. N. (2009). The effect of urban neighborhood disorder on evaluations of the police and courts. *Crime & Delinquency*, *55*(3), 339-362.

St. Louis, E., Saulnier, A., & Walby, K. (2019). Police use of body-worn cameras: Challenges of visibility, procedural justice, and legitimacy. *Surveillance & Society*, *17*(3/4), 305-321. https://doi.org/10.24908/ss.v17i3/4.8649

Sunshine, J., & Tyler, T. R. (2003). The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing. *Law Society Review*, *37*(3), 513–548. https://doi.org/10.1111/1540-5893.3703002

Sykes, R. E., & Brent, E. E. (1983). *Policing, a Social Behaviorist Perspective*. University of Michigan.

Sytsma, V. A., Piza, E. L., Chillar, V. F., & Grossman, L. S. (2021). Measuring procedural justice policy adherence during use of force events: The Body-Worn Camera as a performance monitoring tool. *Criminal Justice Policy Review*, *32*(9), 938–959. https://doi.org/10.1177/08874034211021894

Terpstra, B. L., & van Wijck, P. W. (2023). The Influence of Police Treatment and Decision-making on Perceptions of Procedural Justice: A Field Study. *Journal of Research in Crime and Delinquency*, *60*(3), 344-377. https://doi.org/10.1177/00224278211030968

Terrill, W., Zimmerman, L., & Somers, L. J. (2023). Applying Video-Based Systematic Social Observation to Police Use of Force Encounters: An Assessment of De-Escalation and Escalation within the Context of Proportionality and Incrementalism. *Justice Quarterly*, *40*(7), 1045–1076. https://doi.org/10.1080/07418825.2023.2222819

Tyler, T. R. (2006). *Why people obey the law*? Princeton University Press.

Tyler, T. R., & Huo, Y. J. (2002). *Trust in the law: Encouraging public cooperation with the police and courts*. Russell Sage Foundation.

U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, & Hyland, S. S. (2018). Body-Worn cameras in law enforcement agencies, 2016. In *Bulletin* (Report NCJ 251775). https://bjs.ojp.gov/content/pub/pdf/bwclea16.pdf

Office of Community Oriented Policing Services, Police Executive Research Forum, Miller, L., & Toliver, J. (2014). *Implementing a Body-Worn Camera Program: recommendations and lessons learned*. U.S Department of Justice Office of Community Oriented Policing Services. https://portal.cops.usdoj.gov/resourcecenter/Home.aspx?page=detail&id=COPS-P296

Perkins, M. (2016). Modelling public confidence of the police: How perceptions of the police differ between neighborhoods in a city. *Police Practice and Research, 17*(2), 113-125.

Pickett, J. T., Nix, J., & Roche, S. P. (2018). Testing a social schematic model of police procedural justice. *Social Psychology Quarterly, 81*(2), 97-125.

R: Rescale Variables to a New Range. (n.d.). Retrieved September 20, 2024, from https://search.r-project.org/CRAN/refmans/datawizard/html/rescale.html

USA White House Fact Sheet: Strengthening Community Policing. (2014). White House Office of the Press Secretary Washington, D.C.

Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., Jurgens, D., Jurafsky, D., & Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(25), 6521–6526. https://doi.org/10.1073/pnas.1702413114

Wells, W. (2007). Type of contact and evaluations of police officers: The effects of procedural justice across three types of police–citizen contacts. *Journal of Criminal Justice*, *35*(6), 612-621.

Wentz, E. A., & Schlimgen, K. A. (2012). Citizens' perceptions of police service and police response to community concerns. *Journal of Crime and Justice*, *35*(1), 114-133.

White, M. D., & Fradella, H. F. (2016). *Stop and Frisk: The Use and Abuse of a Controversial Policing Tactic* (6th ed., Vol. 28). NYU Press.

White, D.R. & Ferrandino, J., (2022). Criminal justice students' perceptions amid George Floyd's death and COVID-19. *Journal of Criminal Justice Education*, *33*(1), pp.58-75.

White, M. D., Todak, N., & Gaub, J. E. (2018). Assessing citizen perceptions of body-worn cameras after encounters with police. *Policing*, *40*(4), 689–703. https://doi.org/10.1108/pijpsm-07-2016-0105

Willits, D. W., & Makin, D. A. (2018). Show Me What Happened: Analyzing Use of Force through Analysis of Body-Worn Camera Footage. *Journal of Research in Crime and Delinquency*, *55*(1), 51-77. 10.1177/0022427817701257

Wolfe, S. E., Nix, J., Kaminski, R., & Rojek, J. (2016). Is the effect of procedural justice on police legitimacy invariant? Testing the generality of procedural justice and competing antecedents of legitimacy. *Journal of Quantitative Criminology, 32,* 253-282.

Worden, R. E., & McLean, S. J. (2014). Systematic social observation of the police. In *Oxford University Press eBooks*. https://doi.org/10.1093/oxfordhb/9780199843886.013.026

Worden, R. E., & McLean, S. J. (2017). *Mirage of police reform: Procedural justice and police legitimacy* (p. 268). University of California Press.

Worrall, J. L. (1999). Public perceptions of police efficacy and image: The "fuzziness" of support for the police. *American Journal of Criminal Justice*, *24*, 47-66.

Wu, Y., Sun, I. Y., & Triplett, R. A. (2009). Race, class or neighborhood context: which matters more in measuring satisfaction with police? *Justice Quarterly*, *26*(1), 125-156.

Wu, Y. (2014). Race/ethnicity and perceptions of the police: A comparison of White, Black, Asian and Hispanic Americans. *Policing and Society*, *24*(2), 135-157.

Zimmerman, L. (2023). *Email Communication about BWC Coding* [Personal communication].