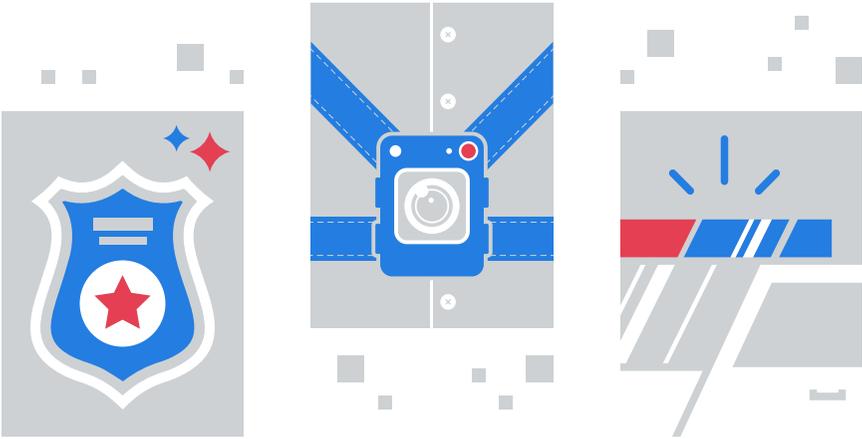


# 16. GREAT EXPECTATIONS, RE-CALIBRATED: EVALUATING DC'S POLICE BODY WORN CAMERA PROGRAM

**David Yokum** is director of The Policy Lab at Brown University and an adjunct associate professor at the Watson Institute for International and Public Affairs. He was previously director of The Lab @ DC in the Government of the District of Columbia, where he remains as senior advisor.

**Anita Ravishankar** is a research scientist at the Metropolitan Police Department in Washington, D.C. and a fellow with The Lab @ DC.



Body worn cameras (BWCs) are purportedly a tool to improve policing and enhance the public legitimacy of a police department. The technology is expensive and, when widely adopted, creates a notable expansion of state surveillance. To what extent do the hoped-for benefits occur, and how does a community balance those benefits against the monetary and privacy costs when deciding whether and how much to invest in a BWC program? What can facilitate informed thinking in this decision, especially amid the emotionally charged backdrop of a nation roiled by a series of high-profile use-of-force incidents, many involving minority residents?

We tell here the story of how Washington, D.C.'s police department handled its BWC program, and in particular, how randomly assigning some but not all officers to wear body cameras and comparing the outcomes helped build the department's capacity to use evidence while generating insights that can inform the national debate.<sup>1</sup>

## ISSUE BACKGROUND

It's March 2015 in the Metropolitan Police Department command center, and the senior police leadership are gathered to make an important decision on the deployment of body-worn cameras in the nation's capital. Since October 2014, the department has been carefully planning the potential deployment of a BWC program to its 3,800-member police force. Pilot work assessed equipment options and settled on Taser International's AXON Body 1; policymakers drafted, debated, and re-drafted regulations over access to video footage; and the city appropriated funds to the program, with a District of Columbia City Council mandate to outfit all officers with BWCs before the end of 2016. The decision point—after

weeks of advance discussion, draft pitches, informal brainstorms over beers, and now a formal proposal in front of those flickering screens—is whether and how to evaluate the program across the 68 square miles of the District of Columbia.

The national movement was remarkable. A series of high-profile, controversial deaths of unarmed African Americans at the hands of police officers, Eric Garner, Michael Brown, Laquan McDonald, Tamir Rice, Walter Scott, sparked a robust national conversation on police accountability and transparency. The role of video footage, either by bystanders and/or dashcam, in bringing many of these incidents to light led directly to a widespread embrace of police body-worn cameras as a technological solution. President Barack Obama proposed a \$75 million, three-year investment to help purchase 50,000 BWCs; the U.S. Department of Justice awarded \$23 million for BWC in 2015 for the initiative.<sup>2,3</sup> A nationwide survey at the time found that 95 percent of large police departments planned to have a BWC program. The central hope was that the watchful lens of a camera would moderate behavior on the streets (either immediately or via a disciplinary and training feedback loop), which in turn would reduce the likelihood that encounters between officers and residents would escalate into violence.

## EVIDENCE AVAILABILITY

Yet the evidence base in early 2015—although promising—was surprisingly thin, especially relative to how rapidly BWC programs were beginning to spread across the country. The primary empirical justification was citation to a single study from Rialto, California.<sup>4</sup> Fifty-four police officers were randomly assigned to wear a camera on some shifts but not others, for a total of 988 shifts over 12 months. Officers were less likely to use force on shifts when wearing a BWC than on shifts when not wearing a camera. Officers were less likely to use force on shifts when wearing a BWC than on shifts when not wearing a camera. No difference in complaints was detected between shift-types. Only an after-the-fact time-series analysis identified a reduction in complaints, a result which was commonly misinterpreted in the media as a finding from the more rigorous evaluation. The authors cautioned against too quickly deciding that BWCs merit expansion, noting in the paper that the results were from “but one experiment.”<sup>5</sup>

Most justification was therefore theoretical, based on evidence from other domains. A wide range of research, dating back to the classic experiments at Hawthorne Works, reported that people may act differently when watched.<sup>6</sup> This literature is not without controversy, but studies have regularly reported—and a common belief is—that we’re more likely to work harder, give to charity, tell the truth, recycle, vote, and so forth if being observed than if alone. If this behavioral effect carried over into the on-street policing environment, then both officers and residents should be more likely to behave in line with community norms when a camera is present. With both parties more closely regulating their behavior, there should be a lower risk of events escalating into violence.

## EVIDENCE USE

This case is about the use of evidence, but in many respects, the most important lessons are about the *generation* of evidence for public consumption. The same evidence may be more or less likely to be used, we believe, as a function of how it is generated and presented in relation to the political decision-making process. But first, it is important to note that the most immediate use of evidence and theory was, of course, in justifying the deployment of the BWC program in the first place. For example, in his testimony before the Committee on the Judiciary, Deputy Mayor for Public Safety and Justice Kevin Donahue relied on the report of an advisory group summarizing the Rialto study and the underlying theory; the Committee Report that informed the City Council's vote to create and fund the program also used this report.<sup>7,8</sup>

Perhaps the most inspiring part of this project was the political desire and will to learn what happened after the initial evidence-informed decision to begin a BWC program. Noting the limitations of prior evidence, Mayor Muriel Bowser, Chief Lanier (and later Chief Peter Newsham), Deputy Mayor Donahue, and City Administrator Rashad Young all recognized the responsibility to evaluate the program as rigorously as possible. A full deployment was inevitable by legislative mandate, so the intention was not to inform a decision to scale; rather, the learning would inform other law enforcement agencies considering BWC programs as well as provide a baseline understanding of what was happening in the District. Such a baseline would provide an evidentiary platform to inform further optimization of the program and, more generally, other activities of the police department.

A decision was made to conduct an experimental evaluation, with each of 2,224 Metropolitan Police Department (MPD) officers randomly assigned—imagine flipping a coin—to either wear or not wear a camera. (The full details of the study, which are a touch more nuanced, are described in a paper under publication in the *Proceedings of the National Academy of Sciences*.) Because of the experimental design with random assignment, outcomes (for example, the likelihood of uses of force or civilian complaints) should be the same on average for the group of officers with BWCs as for the group of officers without cameras, unless the one thing controlled to be different—namely, the assignment of BWCs—*causes* a difference.

The study began in June 2015 with a pilot in two of the seven MPD police districts. This period was to ensure proper implementation of the program and to monitor fidelity to the study protocol. It also, importantly, generated early evidence of potential effect sizes, which researchers used to inform research design options. With immense public pressure to fully deploy the program immediately, the selection of the research design—including how long to run the study and at what level of randomization—was not an easy or simple one. Yet our pilot results indicated that the precision of the effect estimates would increase with the duration of the study. Although there are rule-of-thumb thresholds in the sciences for determining whether an experiment is precise enough to yield meaningful results, the decision of precise to be is ultimately a value judgment. We were able to inform that value judgment—that political decision—by providing power analyses of the minimal detectable effect, assuming the study

lasted three, six, nine, or 12 months. The infusion of this technical information bolstered the political buy-in to wait until the latest legislatively permissible endpoint, December 2016, before distributing BWCs to the group that had not received them initially. Mayor Bowser was even questioned on MSNBC's Morning Joe as to why the department had implemented this delay; she smoothly whipped back that "there was no delay at all," but rather we needed a "control group" to learn how the program was working.<sup>9</sup>

On October 20, 2017, we released a working paper with interactive results on a website dedicated to the study.<sup>10,11</sup> The study found that BWCs had no statistically significant average effects on any of the measured outcomes. To get a feel for the results, imagine groups of officers over the course of a year: a group of 1,000 officers with BWCs was estimated to document 74 more uses of force in a year than a group of 1,000 officers without BWCs, yet the data were also consistent with the real effect of BWCs being anywhere from a decrease of 97 documented uses of force to an increase of 244 documented uses of force per 1,000 officers, per year. A null result obtained for complaints too, with anywhere from a decrease of 24 complaints to an increase of 138 complaints per officers, per year, consistent with the data. We also did not find any downstream courtroom effects in how complaints or cases were adjudicated, although the statistical power was much weaker on this front.

News of the study quickly went global. The study's release coordinated with a front-page article in *The New York Times* and an NPR segment for *All Things Considered*.<sup>12,13</sup> NBC Nightly News interviewed Chief Newsham. Dozens more articles and interviews proliferated.<sup>14,15</sup> Researchers presented at the International Association of Chiefs of Police and fielded numerous phone calls with researchers and leadership teams from police agencies across the country and internationally to educate people about the findings. We also held two community conversations in D.C., to engage with local residents about the study findings and their implications for the city.

We spent an inordinate amount of effort ensuring that no one misunderstood or exaggerated the results or their implications. Indeed, the working paper, website, and associated coverage are perhaps unique in the quantity of pages and airtime given to explaining what the null results do *not* mean, and to highlighting alternative explanations for why we might have found what we did. We noted the risk of spillover, questions about generalizability beyond the D.C. context, limitations of administrative data, the possibility of small or distributional effects outside of what the study methods can capture, and the fact that no survey data was collected about public perceptions of police legitimacy. We also vetted our work, before any public releases, with a series of quality control measures and reviews. All analyses were conducted by two independent statistical teams, for example, which helped avoid coding errors and confirmed a convergence of results. We shared drafts and presentations with peer experts for advance review. Our code and replication data publicly will also be made publicly available.

Our overall recommendation was that we should recalibrate our expectations about the impacts of BWC programs, especially with regard to the possibility of large reductions in the average use of force or complaint rates. We deliberately did not advocate for or against the

adoption of a BWC program. The language was instead carefully crafted to prompt Bayesian thinking, that is challenging pre-existing beliefs based on the evidence. We considered how this study should be incorporated into the broader research and policy debate: the study does not definitively prove that BWCs have no impacts—no single study could accomplish such a feat—but it does recommend updating expectations, recognizing that the likelihood of positive effects on the measured outcomes is lower than previously believed.

It remains a political judgment as to whether a mayor or chief of police wants to implement a BWC program, and there may be justifications beyond seeking large average reductions in use-of-force or complaint rates. Perhaps the main objective is to enhance the public perception of legitimacy (which the study did not measure), or perhaps a jurisdiction would be willing to absorb the considerable financial and privacy costs for the possibility of capturing even a single wrongful act on video (an effect size too small to be practically measured). In the District, for instance, there is ambition to use the video footage for innovative officer training, to which end the city passed a unique legislative provision empowering access to the footage for research purposes.

If other jurisdictions use evidence from the District study to prompt more careful deliberation about these justifications and tradeoffs, then it will be a successful use of evidence, no matter which side of the decision a department ultimately lands on. For example, some people occasionally question the generalizability of the results by noting that the MPD has a unique history of reform and officer training protocols. BWC may be more impactful in police agencies without that history. Maybe indeed. But if people use the study's evidence successfully, it will prompt consideration of whether to prioritize investments in the technology or in implementing reform and training efforts similar to those in the District of Columbia.

## LESSONS

- **Use a pre-analysis plan.** A remarkable feature of this project was the relative lack of after-the-fact attempts to discredit the study's findings on methodological grounds, to attack the independence of the research team (The Lab @ DC is housed out of the Executive Office of the Mayor), or to politically attack the mayor and chief of police for the program results. Discussion was instead, refreshingly, almost entirely about the implications of the evidence and what to do next. We attribute this success to a unique campaign of community engagement, anchored around public registration of a pre-analysis plan. A pre-analysis plan documents the questions a study will ask and the methodologies it will use in answering those questions—all written down before looking at or even collecting the data. Different methodological choices can generate different answers, and research has uncovered that scientists, consciously or more often unconsciously, have a bias toward choices that generate a desired result (so-

called “*p*-hacking”). A related bias is the tendency to tell stories after results are known, as if the result were theoretically expected from the onset (so-called “HARKing,” or hypothesizing-after-results-known).<sup>16</sup> Both biases are short-circuited if methodological choices are not made in advance of conducting data analysis.

- ***Lean into the reality that science involves politics.*** The pre-analysis plan is also a unique vehicle for enhancing the *political* integrity of a project.<sup>17</sup> Science involves value judgments: how large of an effect size, measured at what precision, justifies a decision to take one action over another? Objective methodologies can provide reliable statistics, but weighing the risks and tradeoffs of any given intervention is always a political act. Lean into this reality. In particular, we used the drafting of the pre-analysis plan to work closely with key decision-makers to facilitate advance thinking about what the most important questions are, how much effort should be expended on answers, and what methods are most suitable. We publicly registered a draft pre-analysis plan on the Open Science Framework in October 2016 to facilitate community feedback.<sup>18</sup> Of course, not many people casually browse the Open Science Framework archives, so a critical step was hitting the streets: between October 2016 and June 2017, we held 13 public events to discuss the pre-analysis plan, ranging from professional audiences, to non-profit and advocacy groups, to everyday audiences gathered at libraries and schools. We circulated the pre-analysis plan widely and invited comments. We worked closely with press offices to ensure that officials understood and could confidently discuss the details when the opportunity surfaced—from council hearings to *Morning Joe*. Our proactive and open approach bolstered trust in the research. When the results came out, there was not the usual arguing over methods because the methodology, whatever its lingering shortcomings, had already been extensively discussed and accepted. When people scrutinized our independence, we could point to the publicly registered pre-analysis plan and explain again how the transparency curbed against *p*-hacking and HARKing.
- ***Co-production of research increased the likelihood of use.*** From the onset, we embedded closely with MPD. We joined internal meetings, conducted ride-alongs, learned to use the equipment, and visited the training academy. The department even hired one researcher as a staff member. And likewise, the department was authentically part of the research enterprise: Chief of Staff Matthew Bromeland, Commander Ralph Ennis, Special Assistant Heidi Fieselmann, and BWC Program Coordinator Derek Meeks worked closely with us to operationalize the field experiment and joined community discussions. This integrated approach empowered us to design a rigorous methodology meshed into the operational realities of what it takes to deploy thousands of pieces of physical equipment across an entire police force. Perhaps the most heartening outcome is how the project elevated the overall capacity of the police department to use and generate evidence into the future. We mapped and documented the administrative data of the department, a resource which is now being leveraged in new projects. MPD also created three permanent positions—a research scientist, a data scientist, and a management analyst—to boost its internal capacity to do research and

evaluation work, and has since initiated a variety of projects to improve recruiting, training, and violence reduction efforts. Terms like “counterfactual,” “confidence interval,” and “effect size” are commonly used in the hallways and meeting agendas, with the question of “how are we going to know if this works?” a common follow-up to any new proposal. ■

- 1 David Yokum, Anita Ravishankar, and Alex Coppock. *Proceedings of the National Academy of Sciences*, forthcoming 2019.
- 2 Andrea Peterson, “President Obama wants to spend \$75 million to buy police bodycam.” *The Washington Post*, December 1, 2014.
- 3 U.S. Department of Justice. “Justice Department Awards over \$23 Million in Funding for Body Worn Camera Pilot Program to Support Law Enforcement Agencies in 32 States.” September 21, 2015. Available at: <https://www.justice.gov/opa/pr/justice-department-awards-over-23-million-funding-body-worn-camera-pilot-program-support-law>.
- 4 Barak Ariel et al. “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, 2015.
- 5 Ibid.
- 6 For a recent review of research on the Hawthorne effect, see: Jim McCambridge, John Witton, and Diana R. Elbourne. “Systematic Review of the Hawthorne Effect: New Concepts Are Needed to Study Research Participation Effects.” *Journal of clinical epidemiology*, 67(3): 267-277, 2014.
- 7 B21-351, the “Body-Worn Camera Program Regulations Amendment Act of 2015”, B21-356, the “Public Access to Body-Worn Camera Video Amendment Act of 2015”, and PR 21-327, the “Metropolitan Police Department Body-Worn Cameras Approval Resolution of 2015”: *Hearings Before the Committee on the Judiciary*, 2015 Leg., (Washington, DC, 2015), statement of Kevin Donahue, Deputy City Administrator, Deputy Mayor for Public Safety and Justice. Available at: [https://mpdc.dc.gov/sites/default/files/dc/sites/mpdc/publication/attachments/FINAL%20Donahue%20Testimony\\_10%2021%2015.pdf](https://mpdc.dc.gov/sites/default/files/dc/sites/mpdc/publication/attachments/FINAL%20Donahue%20Testimony_10%2021%2015.pdf).
- 8 Council of the District of Columbia, Committee on the Judiciary, *Report on Bill 21-0351, the “Body-Worn Camera Program Amendment Act of 2015”* Council Period 21, 2015. Available at: <http://lirms.dccouncil.us/Download/34469/B21-0351-CommitteeReport1.pdf>.
- 9 Muriel Bowser, interview by Joe Scarborough and Mike Brzezinski, *Morning Joe*, MSNBC, July 14, 2016.
- 10 David Yokum, Anita Ravishankar, and Alexander Coppock. “Evaluating the Effects of Police Body-Worn Cameras.” Working Paper, October 2017. Available at: [https://bwc.thelab.dc.gov/TheLabDC\\_MPD\\_BWC\\_Working\\_Paper\\_10.20.17.pdf](https://bwc.thelab.dc.gov/TheLabDC_MPD_BWC_Working_Paper_10.20.17.pdf).
- 11 The Lab @ DC, “Randomized Controlled Trial of the Metropolitan Police Department Body-Worn Camera Program,” 2017. Available at: <https://bwc.thelab.dc.gov/#home>.
- 12 Amanda Ripley, “A Big Test of Police Body Cameras Defies Expectations,” *The New York Times*, October 20, 2017.
- 13 Nell Greenfield, “Body Cam Study Shows No Effect on Police Use of Force or Citizen Complaints,” *All Things Considered*, National Public Radio, October 20, 2017.
- 14 Peter Herman. “Police Officers with Body Cameras Are as Likely to Use Force as Those Who Don’t Have Them.” *The Washington Post*, October 20, 2017.
- 15 Amanda Ripley. “New Study Finds Little Change in D.C. Officers’ Behavior with Body Cameras.” *The Seattle Time*, October 20, 2017.
- 16 For an accessible and interactive introduction to issues such as p-hacking and HARKing, see: Christie Aschwanden. “Science Isn’t Broken: It’s Just a Hell of a Lot Harder Than We Give It Credit For.” *FiveThirtyEight*, August 19, 2015. Available at: <https://fivethirtyeight.com/features/science-isnt-broken/#part1>.
- 17 David Yokum. “Psychology, Open Science, and Government: The Opportunity.” *Association for Psychological Science Observer*, 29(4), 2016. Available at: <https://www.psychologicalscience.org/observer/psychology-open-science-and-government-the-opportunity>.
- 18 David Yokum. “Pre-Analysis Plan: A Randomized Controlled Trial of the Police Body-Worn Camera Program in the District of Columbia.” October 7, 2016. Available at: <https://osf.io/yjyng/>.