What Counts as Police Violence? A Case Study of Data in the CATO Institute’s Police Misconduct Reporting Project

Anna Feigenbaum & Dan G.B. Weissmann
Bournemouth University

ABSTRACT
Background This article presents a case study about the role of data in the CATO Institute’s Police Misconduct Reporting Project and reflects on what constitutes police violence.
Analysis Augmenting this data aggregation work, the article turns to additional data projects focused on recording police crime and misconduct to gather a broader understanding of incidents of police violence beyond acts that cause death.
Conclusion and implications It is only when we look at data on acts of violence that occur when an officer is on duty and off-duty, with or without a firearm, that a clearer sense of the traumatic cycle of policing can be understood. This way of looking at police data requires both broader practices of “copwatching,” as well as a broader definition of what counts as violence.

Keywords Data; Police violence; Spillover violence; Toxic masculinity; Copwatching

RÉSUMÉ
Contexte Cet article présente une étude de cas sur le rôle des données dans le Police Misconduct Reporting Project (Projet sur le recensement des inconduites policières) du CATO Institute et offre une réflexion sur ce qu’est la violence policière.
Analyse L’article complète cette étude de cas en examinant d’autres projets semblables portant sur les inconduites et crimes policiers afin d’en arriver à une meilleure compréhension d’instances de violence policière au-delà d’actions causant la mort.
Conclusion et implications C’est seulement quand on examine des données sur les actes violents ayant lieu quand un policier est de service ou non, avec ou sans arme à feu, qu’on peut mieux comprendre le cycle traumatique relatif au travail policier. Cette manière de considérer les données sur la police requiert à la fois des pratiques plus englobantes de surveillance des policiers et une définition plus large de ce qu’est la violence.
Mots clés Données; Violence policière; Débordement de violence; Masculinité toxique; Surveillance de policiers

Anna Feigenbaum is Principal Academic in Digital Storytelling in the Department of Journalism, English and Communication at Bournemouth University, Talbot Campus, UK. Email: afeigenbaum@bournemouth.ac.uk. Dan G.B. Weissmann is a PhD Researcher in Political Communication and Associate Lecturer of International Relations in the Department of Humanities and Law, Bournemouth University, Talbot Campus, UK. Email: dweissmann@bournemouth.ac.uk.

Introduction
The police shooting of unarmed black teenager Michael Brown on August 9, 2014, ignited uprisings in Ferguson, Missouri. As these protests made national and international headlines, questions of police violence, particularly against African-American communities, came to the fore in the United States. Activists, advocacy groups, and journalists used this moment to expose the poor state of official record-keeping on officer-involved shootings. Prior to Ferguson, the operational system in place for recording police killings was unsystematic and voluntary. Individual law enforcement agencies—of which there are more than 18,000 in the United States—could choose whether or not to submit their yearly statistics on “justifiable homicides,” which were defined as “the killing of a felon in the line of duty.” (Swaine, Laughland, Larney, & McCarthy, 2015). Between 2005 and 2012, only 1,100 agencies reported justifiable homicides (Swaine, Laughland, Larney, & McCarthy, 2015). It quickly became clear the police had a data problem.

Making this data problem visible on an international scale, major news projects began to aggregate, verify, and analyze incidents of police killings, distributing this information to the public through interactive graphics and data visualizations. These projects by the Guardian (The Counted), the Washington Post (Fatal Force), and earlier efforts began by the Fatal Encounters project, marked a new era in police accountability. Their innovation and amplification was tied to the rise of collaborative social media technologies and platforms. Wiki-style websites, Twitter reports, encrypted emails, and local news stories that could be shared with the click of a button, together with more traditional forms of reporting, made these data journalism projects possible.

From copwatching to data aggregation
While these data practices are new, the recording and monitoring of the use of force by police has a long history in “copwatching.” As defined by legal scholar Jocelyn Simonson (2016 p. 393), copwatching is “groups of local residents who wear uniforms, carry visible recording devices, patrol neighbourhoods, and film police-citizen interactions in an effort to hold police departments accountable to the populations they police.” In the United States, formalized copwatching dates back to at least the 1960s, with Black Panther patrols and community police watching established during the Civil Rights and Black Power era (Simonson, 2016).

As digital technologies proliferated, making video cameras more mobile and accessible, practices of filming the police expanded. With the rise of social media, activist practices around documenting the police have taken on new levels of sophistication, including using secure and anonymous smartphone apps and live-streaming technology (Feigenbaum & McCurdy, 2018). Shared over Facebook and Twitter, these bite-size broadcast videos can circulate around the world within hours of an incident (Bennett & Segerberg, 2013; Gerbaudo, 2012). Such innovations in copwatching practices also responded to police intimidation and the destruction of filmed evidenced (Bock, 2016). Today, the rise of police body-worn cameras serves as both a response to the call for more accountability and an attempt to put the narrative power and perspective (or gaze) back in the hands of the police (Brucato, 2015, Feigenbaum & McCurdy, 2018).
While there is no scope in this article to explore the trajectories between these data journalism projects and community-based copwatching, two important observations can be made:

Practices of counting, or aggregating incidents of police violence, both rely on and go beyond more traditional notions of copwatching, such as the recording and monitoring of localized police violence.

These practices of counting, when done by civilians, form part of a broader tactical repertoire of copwatching that leverages the power of statistics and quantitative, evidence-based reporting.

These two observations are not put forward to argue that data practices are superior to more traditional forms of copwatching but rather to draw attention to the broader trajectories and assemblages of police accountability projects, bridging the analogue and digital, the datafied and street-based reporting. Moreover, as will be argued in relation to the case study on police misconduct monitoring:

These varied tactics of copwatching are crucial for understanding police violence beyond officer-involved shootings, expanding our definitions to include a wider range of use-of-force incidents.

Case study rationale
Diverging from other work in this area that is focused around the question of accountability (Brucato, 2015; Gray, Lämmerhirt, & Bounegru, 2014), this article works from Steve Martinot’s (2013) argument that there is a structural culture of police impunity and a regimented demand for obedience that justifies the use of force. It is only when we look at violence on duty and off-duty, with or without firearms, that a clearer sense of the traumatic cycle of policing can be understood. This way of looking at police data requires both broader practices of copwatching, as well as a broader definition of “what counts” as violence. Drawing from Leanor Johnson, Michael Todd, and Ganga Subramanian (2005), discussions of spillover violence, and Tracy Karner’s (1996) work on toxic masculinities, this article argues that aggregations of data on police use of force must look not only at the question of if someone is killed but also at how violence is embedded and manifested in the full cycle of police lives—from the streets to the home.

Stinson’s police crimes project
While projects such as Fatal Encounters, Fatal Force, and The Counted highlighted the “epidemic” of killings by police, it is not only more data on deaths that advocates, academics, and policymakers have called for. Other types of police crime and misconduct, from drug dealing and embezzlement to domestic assault and child molestation, also suffer from what Philip M. Stinson, John Liederbach, Steven L. Brewer Jr., and Natalie E. Todak (2016, p. 188) call “an absence of suitable data.” As of the time of writing, there is no organized system for the collection, monitoring, or analysis of data on police crime:

The traditional sources of data and methods of study, whether official statistics, self-report surveys, or direct observations, either do not exist in any usable format or are ill-equipped to identify, count, or provide the basis for empirical analyses of instances in which police perpetrate crimes themselves. (p. 26)
Similar to those journalists working on Fatal Encounters, Fatal Force, and The Counted, Stinson’s (2016) team decided to approach the news as a primary data source. They used Google alerts to identify, print, log, scan, index, and digitally archive thousands of articles on police crime. Yet, taking a social sciences approach, they recognized both police data and news data as questionable sources.

Stinson et al. (2016) argue that the amount of media attention given to police crimes can be the result of the intrinsical newsworthiness of such stories. This idea of police crime as newsworthy can be elaborated through the more nuanced language of news values. News values refers to a body of research that has shown the presence of a shared set of features that is valued by the news industry. These features guide the process of storytelling and publishing, turning an event or incident into news (Bednarek & Caple, 2014; Galtung & Ruge, 1965). Police crime can be seen as fulfilling a number of these “news values,” including tipping past a threshold of acceptable behaviour, the negativity of content, involving people in power, and personification or stories on individual people (Galtung & Ruge, 1965). Incidents of police crime also often involve sensationalism (Östgaard, 1965), have a novelty value (Gans, 2004), relate to deviance and thus have social significance ( Shoemaker, Danielian, & Brendlinger, 1991). Where there is a perpetrator and a victim, these incidents also often have an in-built conflict, which is an essential feature of narrative storytelling (Johnson-Cartee, 2005). These limitations or “data distortions” (Feigenbaum, Thorsen, Weissmann, & Demirkol, 2016) were important considerations in this case study’s approach to working with data from the CATO Police Misconduct Reporting Project. The problems of data distortion will be revisited in the article’s conclusion.

The CATO Police Misconduct Reporting Project
Also attempting to fill this void in data collection around police misconduct, in 2009 independent researcher David Packman started the Police Misconduct Reporting Project. The libertarian CATO Institute took over this monitoring project in 2012 but discontinued the project in 2017 due to a lack of resources. This reporting project used Web crawlers to scrape news stories and supplements this with research done by interns. Each week CATO publishes a round-up of news reports on police misconduct. These go out on its project website, as well as social media. The public nature of sharing these scraped data stories is designed to be a kind of live stream of accountability. Mobilizing a logic similar to that found in copwatching, data in the CATO project is used performatively, rather than as a symbol of the real. In other words, the value of data for the CATO project is not found in its authority or truthfulness but in its strategic ability to keep the police under public scrutiny through the act of making data sources openly accessible and available. By presenting data as micro-stories, rather than complex infographics or full databases, the data becomes easily shareable, enhancing its storytelling quality to connect with readers. Yet, what was eye-catching in the CATO data was not its social media shareable form but rather its potential to be wrangled.

Methodology for using CATO’s dataset
Because the CATO Police Misconduct Monitoring Project used a structured dataset to derive its weekly updates, it was clear its format would make it possible to scrape,
geocode, and visualize the data to produce an augmented dataset. Initially the aim was only to generate another “media map” that visually represented the coded information in an easy to use “drill-down” interactive for sharing information and embedding in news stories and blogs. Similar to these other “counting” projects, the media maps of mass tear gasings responded to the lack of official data by turning media reports into data sources to provide insights where no official data were available, what Christine Borgman (2015, p. 166) calls an area that is “data poor.” Following lesson number one from data journalism, “Always ask before you scrape,” we contacted Jonathan Blanks who ran the CATO project and asked if he would be willing to share his dataset. Pleased with the opportunity to have more eyes on the dataset, he not only provided us with a relatively clean set but also a link to the project’s own methodology.

CATO’s methodology
The CATO dataset itself consists of 2,001 incidents of police violence during the year 2015, recorded via local and national news sites in the U.S. The news sites span a variety of outlets, from well-established newspapers such as the Miami Herald (Ovalle, 2015) to small, rural news outlets such as the Victor Valley News (2015) in Victorville, California. It is important to keep in mind here that the variety of news outlets also represent a variety of journalistic practices, such as professional journalism, volunteer journalism, and citizen journalism. Since there is a lack of recording violent police encounters on an institutional level in the U.S., the CATO Institute had to rely on local and national news sources to create the dataset (Stinson et al., 2016). These were collected using a Web crawler, as well as various individuals who collect their own data, and individuals on Twitter who pointed the CATO team toward relevant news stories. For scope, the dataset only included full-time, active duty, sworn officers. Using only sworn officers allows for consistency in the context of legal education, as the law applies differently to part-time officers and non-sworn officers, for example, bailiffs. Also, the CATO project is interested in researching accountability for police crime and misconduct respectively in the context of whether police are disproportionately found not guilty of charges, or do not face charges. Therefore, CATO’s own coding is a reflection of this focus on legal education.

While it was not the original intention to analyze this data through a gendered lens, during the process of preparing the data for visualization, the amount of gendered and sexualized violent incidents documented became apparent. This resonated with previous findings around vulnerability, masculinity, and trauma in representations of police use of force (Feigenbaum & Weissmann, 2016). In order to draw out occurrences of such gendered and sexualized crime, it was necessary to adapt and expand the codes used by the CATO project. This meant moving away from purely legal terminology to incorporate additional codes that could capture these aspects of misconduct and violence, expanding the notion of what counts as police violence.

Data analysis
Coding the dataset for the purpose of this article required cleaning the individual entries for spelling and formatting errors using OpenRefine (2010). OpenRefine was also used to cluster and eliminate duplicates, thereby avoiding distortion. Incidents marked
“updates” were removed from the sample except in cases where a) there was more than one charge or accusation, b) there was an administrative and criminal step to the process (i.e., someone resigned and was later charged), c) the data were marked update, but the original post was not in the sample, d) additional or different people were involved in the update. In other words, updates were deleted when there was no change in the officer involved and the type of misconduct. These changes gave the dataset the purpose of “incident reporting” rather than “incident tracking.”

Importantly, this meant the dataset should be seen as a snapshot in time, not as a longitudinal study of police misconduct. Meaning that this study does not focus on what happens to the individual officer or case, hence the removal of officer names from the dataset. Additionally, the dataset is in no way a representative sample as it only contains news reporting and does not rely on official databases, as no such databases existed at the time of writing for the reasons discussed above.

Seeing police violence as a structural issue that goes beyond on-duty police officers, as Steve Martinot (2013) suggests, called for extending the coding scheme and including the concept of spillover violence (Johnson et al., 2005) to not only include on duty but also off duty incidences. Spillover violence is violent behaviour towards a spouse or children through work related conditions such as exposure to violence on the job, burnout in the job, and substance abuse (Johnson et al., 2005). In short, it is bringing the occupational violence and stress on the job back into the family home (Beehr, Johnson, & Nieva, 1995).

To operationalize the concept of spillover violence, a coding frame was developed to designate types of sexualized and gendered violence, as well as violence against domestic partners/spouses and minors. Additionally, a code for firearm was added to identify instances of a firearm in a domestic setting as highlighted by Johnson et al. (2005) and Philip Stinson et al. (2013). Drawing in part on Stinson’s previous work, the codes used for this dataset were: sexual misconduct, sexual assault, soliciting, sex with a minor, child molestation, child pornography, domestic assault, assault, assault minor, and vehicle assault. Since not all codes clearly identify the age of the victim, an additional column was created to include a y/n function for the involvement of a minor. The same method was used to create a y/n function for the involvement of a firearm. Using these codes, an initial pilot set of 100 entries were coded by two coders (the authors) to test for intercoder reliability before coding the full dataset.

Findings and discussion

Sexualized/gendered violence and toxic masculinity
The dataset quickly revealed the sexualized and gendered dimensions of police violence, not only on duty to female and minors as suspects or witnesses but also, and more dominantly, domestic violence off-duty toward spouses and children. Taking all sexualized violence categories together, it is present in 256 out of 2,001 incidences, making it the most common reported incidence among police officers on duty and off-duty. Additionally, the dataset showed domestic assault charges almost equal to the on-duty assault charges, with 171 incidences of domestic assault off-duty and 172 incidences of assault on duty.
While it should be remembered that this sample only represents reported incidents and only provides a snapshot over a short period of time, as a descriptive finding it supports social science research into gendered and sexulized abuse and spillover violence in policing. In particular, it speaks to the cultures of “toxic masculinity” that emerge out of the hypermasculine and violence-oriented environment of contemporary policing. The concept of “toxic masculinity” emerged out of interviews conducted with Vietnam veterans. Looking at the effects of wartime experiences on health and behaviour at home, Karner’s (1996) groundbreaking work on post-traumatic stress disorder (PTSD) found that:

Next to the adrenaline rush of combat and the challenges of survival, the normative structures of job, family, church, and community seemed, for many, banal and superficial … [These men] eventually stopped playing those roles and began to expend more and more energy on activities that I refer to as “toxic masculinity,” such as excessive drinking, almost compulsive fighting and violent competition with other men or male authority figures, dangerous thrill seeking, and reliving or reenacting combat behavior in their stateside environments. (pp. 76–77)

Drawing on Karner’s (1996) work, and that of other scholars looking at masculinity and cultures of violence, Leigh Goodmark (2015) argues that violence and the denigration of women are endemic in policing. Authoritarian training styles and repeated training exercises in the use of coercive force are also linked to what Johnson et al. (2005) call a spillover of violence at home.

While empirical data are limited due to the failure of law enforcement to monitor its crimes or provide public figures, a set of statistics compiled by the Feminist Majority Foundation’s (2013) National Center for Women and Policing in the early 2000s suggested that the rate of domestic abuse in policing is up to four times higher than the national average, with 40 percent of families experiencing abuse of some kind. More recent data collected by Stinson’s (2013) team, and supported by the data from the Police Misconduct Project, suggest that at least one in every five cases of recorded police misconduct are of domestic and family abuse. Even if these data are distorted, all of these small-scale studies, taken together, repeatedly suggest that police have much higher rates of domestic and child abuse in the family than the general population.

On-duty/off-duty and spillover violence
The vast majority of these domestic abuse incidents happen off-duty. But toxic masculinity and its violence is not only spilling over into the home. In the CATO dataset, over 40 percent of recorded reports of police misconduct in 2015 occurred off-duty. Drunk driving incidents almost all happen exclusively off-duty, with 172 incidents off-duty and only two incidents on duty. The same is the case for domestic assault, where 156 incidents happened off-duty and only nine incidents on duty. However, incidents of non-domestic assault are almost split evenly, with 77 incidents off-duty and 88 incidents on duty. Furthermore, incidents of misconduct reported while an officer was on duty are almost exclusively procedural violations, with 179 on duty and 14 off-duty. The use of excessive force is also recorded almost exclusively on duty, with 162 on-
duty and only seven off-duty reports. Finally, theft is recorded as one of the incidents that is reported predominantly in an on-duty context, with 144 incidents on duty and 30 off-duty.

While incidents of theft, excessive force, and non-domestic assault are primarily reported while an officer was on duty, the data shows that domestic assault occurred mainly off-duty. This suggests what Johnson et al. (2005) refer to as spillover violence, where police officers behave as though they are still on duty at home, or as Johnson et al. (2005) put it, they act like “police officers at home” (p. 7). The term spillover captures this “spilling over of violence,” looking at how stresses of the job get carried over into off-duty life at home.

Spillover violence is reflected most explicitly in reports of domestic violence, as the CATO dataset and Johnson et al. (2005) point out. In addition, spillover can also be seen in the use and misuse of alcohol by off-duty officers. The high number of off-duty drunk driving incidents reported in the dataset attests to how stresses of the job can spillover into the rest of an officers’ life. The Stinson et al. (2013) study of police misconduct incidents found that drunk driving was a frequent type of legal violation that officers are not appropriately held accountable for, while Johnson et al. (2005) link alcohol abuse directly to incidents of domestic violence.

Conclusion
Building on an intervention made by Ben Brucato (2015) in relation to copwatching and the new transparency, acts of data aggregation function as a mirror, reflecting back glimpses of truth some may not wish to see. Yet crucially, it is not the counting alone that stands to transform police cultures of violence. Better data on police killings—or even on police misconduct more broadly—do not guarantee better understandings of how and why violence functions with impunity in policing. For this richer understanding of cultures of violence, data projects must be integrated and must work toward broader understandings of what counts as police violence. Incidents of violence conducted while on duty cannot be isolated from the life cycle of violence more broadly as it spills over into off-duty life. Rather, we need a broader view of violence that accounts for abuse and harms as they occur on the street, at the workplace, or in the home. To better account for this, further research might explore how cultural institutions beyond the police academies and law enforcement agencies contribute to these cultures of toxic masculinity and spillover violence. From promotional police ads to use-of-force workshops and the for-profit sales of ever-more technological arsenals, police culture is marketed and reproduced through a broader network of institutions and organizations that may be thought of as part of a use-of-force industry (Feigenbaum & Weissmann, 2016).

In addition, as Borgman (2015) argues, it is crucial that we not only share data with each other but also that we share reflections on our experiences around the processes of generating, cleaning, and analyzing data. For example, it is crucial that we disclose the limitations of our datasets, its inherent biases, and how processes of data collection can cause data distortions. For example, in the CATO dataset, a large number of incidents are from Florida. While it might at first appear this is because of its population has a propensity toward the use of force, Jonathan Blanks (interview...
with authors) suggests that it may instead be due to stricter rules for reporting and sharing data on police misconduct in the state.

Finally, it is necessary to move beyond the idea that simply opening data creates accountability, instead there must also be a shared knowledge infrastructure that allows us to look at problems from multiple perspectives at once. Building infrastructures requires conceptually rich modes of coding, attunement, and honesty about data distortions, and a value system that recognizes the relationship between traditional and data-driven reporting as well as analogue and digital modes of copwatching. As these different forms of data collection and analysis come together, they can reveal broader understandings about police violence and its impacts on civilians, officers, and their families.

**Interview**


**References**


