Officer-Involved: The Media Language of Police Killings*

Jonathan Moreno-Medina ⓡ Aurelie Ouss ⓡ Patrick Bayer ⓡ Bocar Ba

July 5, 2022

Abstract

This paper studies the language used in television news broadcasts to describe police killings in the United States from 2013-19. We begin by documenting that the media is significantly more likely to use several language structures - e.g., passive voice, nominalization, intransitive verbs - that obfuscate responsibility for police killings compared to civilian homicides. We next use an online experiment to test whether these language differences matter. Participants are less likely to hold a police officer morally responsible for a killing and to demand penalties after reading a story that uses obfuscatory language. In the experiment, the language used in the story matters more when the decedent is not reported to be armed, prompting a final research question: is media obfuscation more common in high leverage circumstances, when the public might be more inclined to judge the police harshly? Returning to the news data, we find that news broadcasts are indeed especially likely to use obfuscatory language structures when the decedent was unarmed or when body camera video is available. Through this important case study, our paper highlights the importance of incorporating the semantic structure of language, in addition to the amount and slant of coverage, in analyses of how the media shapes perceptions.

JEL Codes: K14, K42, L82

*We are grateful to Mackenzie Alston, Kareem Haggag, Anjelica Hendricks, Justin Holz, Dean Knox, Rachel Mariman, Arnaud Philippe, John Rappaport, Martin Salzmann and Jennifer Tamas for helpful comments. We thank Jasmine Carter, JoonYup Park, Caroline Milgram and Adam Soliman, for providing excellent research assistance. The author order has been randomized and recorded on the AEA Author Randomization Tool, with Confirmation ID 'LPFDU1546yLy'. The randomization of order is indicated by the symbol ⓡ. Moreno-Medina: Economics Department, Alvarez School of Business, University of Texas-San Antonio, Ouss: Department of Criminology, University of Pennsylvania, Bayer and Ba: Department of Economics, Duke University.

Keywords: Police killings, homicide, media, news coverage, semantics

> As officers contacted the suspect an OIS [Officer Involved Shooting] occurred, one of the officer's rounds penetrated a wall that was behind the suspect. Beyond that wall was a dressing room. Officers searched the dressing room and found a 14 year old female victim who was struck by gunfire.
>
> *Tweet from Los Angeles Police Department Media Relations following police killings of Valentina Orellana-Peralta and Daniel Elena Lopez, 2021*

> Things [...] can indeed be defended, but only by arguments which are too brutal for most people to face, and which do not square with the professed aims of political parties. Thus political language has to consist largely of euphemism, question-begging and sheer cloudy vagueness.
>
> *George Orwell, 1946*

1 Introduction

The language used by the news media to describe an event matters. The choice of language affects how people understand and imagine what happened, their perceptions about causality and moral responsibility, and, ultimately, their broader beliefs and judgments about the world around them (Pinker, 2007). Given its central role in society, many dimensions of media coverage have received considerable academic attention, including what events are covered in the first place (Eisensee & Strömberg, 2007; Enikolopov *et al.*, 2011) and the words that are used to describe them - e.g., political slant/bias or gendered language (Gentzkow & Shapiro, 2010; Chiang & Knight, 2011; Martin & Yurukoglu, 2017; Jakiela & Ozier, 2018; Gay *et al.*, 2018).

In this paper, we examine another critical aspect of media language – semantics – which is broadly concerned with how the structure of language affects its understood meaning. Specifically, we study the use of particular sentence structures, such as the active versus passive voice or the inclusion versus omission of a subject, which systematically work to either clarify or obfuscate the actor and/or actions taken during an incident. We do this in the context of the media coverage of police killings of civilians. Over 1,000 individuals are killed by the police annually in the US, accounting for about four percent of all homicides.[1] Media watchdogs have called attention to the tendency for news reports and police department press releases to describe police killings using language structures specifically designed to diminish the central, active role of police officers in the killing. Journalist Radley Balko has coined the term "exonerative tense" for these language structures to highlight their apparent aim of dampening negative judgments about the appropriateness of the officer's actions.[2]

Our paper proceeds in three steps. First, using data on the universe of American television news broadcasts (both local and national) from 2013-19, we examine whether there is greater use of obfuscatory language structures in coverage of police killings versus homicide reporting in general. We then use an online experiment to causally test whether obfuscatory language matters for how people understand a news story about a police killing, how they assign agency and responsibility, and their broader perspectives on the issue. Finally, we return to the news data to examine whether obfuscatory language is used more frequently by the media in circumstances when our experiment suggests it would have the greatest impact on a viewer's perception.

[1]https://www.cdc.gov/nchs/fastats/homicide.htm

[2]See for example Balko (2014) or Blachor (2020) for such discussions in popular press.

To characterize how language structure affects clarity/obfuscation about actors and actions, we draw heavily on the linguistics literature (Toolan, 2013; Pinker, 2007). We capture four dimensions of potential obfuscation relative to an active sentence in which a police officer is the subject (`‘police officer killed man’`). The first is the use of the *passive voice* (`‘man was killed by police officer’`), which pushes the role of the police officer to the background of the sentence, potentially decreasing its salience to the reader. The second is a further transformation of the sentence to remove any reference to the police as the cause of the killing (`‘man was killed’`). We refer to this structure as: *no agent*. A third obfuscatory structure is the use of *nominalization*, which involves transforming the action of the police killing into a noun (`‘deadly officer-involved shooting’`). The final dimension we consider is the use of the intransitive – e.g., transforming the transitive verb ‘kill’, which requires an agent who generates the action, into the *intransitive* ‘die’, which does not require a cause (`‘man dies [in shooting]’`).

Our primary data set combines text captions from the universe of American television news broadcasts from 2013-19 (covering national and local stations) with data on the universes of: (i) police killings of civilians drawn from the Mapping Police Violence (MPV) database; and (ii) civilian gun homicides drawn from the Gun Violence Archive (GVA). We capture the four dimensions of obfuscation described above using recently-developed Natural Language Processing algorithms for co-reference (i.e. finding all expressions that refer to the same actors) and semantic role labeling tasks (Lee *et al.*, 2018; Shi & Lin, 2019).[3] In order to benchmark the coverage of police shootings, we compare it with the coverage of civilian homicide shootings. We restrict the comparison to homicides in which a suspect has been identified somewhere in the story to ensure that the media's potential language choices are comparable.[4] We conduct a number of robustness checks (such as dropping the sentences in which the suspect's name appears) to ensure that the requirement that the suspect's name appears in the story does not bias the results. Our sample includes a total of roughly 6,000 police killings, 8,000 non-police homicides, 200,000 stories, and 470,000 sentences that describe the killings.

Our main findings are as follows: we find higher obfuscation in the description of police killings than in other homicides across all four dimensions of obfuscation. Overall, there is some obfuscation in 35.6 percent of stories about police killings, compared to 28.8 percent

[3] We explain the essence of these tasks in Section 3.3 and describe our implementation in further detail on Appendix A.2 and A.3

[4] As we explain in more detail below, if the shooter is not known (as is often the case in the immediate aftermath of an incident), it is more natural for the media to use the passive voice (*a victim was killed*) versus the active (*unknown person kills man*) in reporting the event.

for civilian homicides, a 25 percent increase. Estimated effects are even greater when we include media market or station fixed effects and, importantly, when we restrict attention to the first sentence of the story. For first sentences, obfuscation is 40 percent greater for coverage of police killings compared to civilian homicides, suggesting that the media employs more obfuscation in the most salient part of the story.

We then proceed to test whether obfuscatory language changes how people understand and process the information in a news story. Although such an effect has been hypothesized in the linguistics literature (Toolan, 2013), there is scant experimental or empirical evidence to date. We conducted an online experiment on Prolific with 2,402 participants to evaluate how sentence structure affects responses to a news story about a police killing. We focus on three main outcomes: judgment about the police officer's moral responsibility in the incident, demand for penalties for the officer (departmental and legal), and financial support for an organization supporting police reform. Our baseline is an active sentence structure, and we compare responses depending on the degree of obfuscation: passive voice, passive with no agent (using "officer-involved" instead), and the use of an intransitive verb (again, with "officer-involved"). We test two main hypotheses: i) obfuscatory sentence structures decrease the perceived moral responsability of the police, the demand for penalties, and support for police reform; and ii) the degree of obfuscation (intransitive being the highest, passive the lowest) is reflected in the order of effect sizes. We also vary each treatment arm by whether the person killed by the police is described as holding a weapon or not. We registered all of our hypotheses before conducting the experiment, including both a direct effect of the reported presence of a weapon (decreases negative judgment against the officer) and an interaction effect (language structure matters less when the victim is reported to have a weapon).

In line with our hypotheses, we find strong evidence that obfuscatory language matters, and that, directionally, the effect sizes are increasing in the degree of obfuscation. As predicted, we also find that the effects are larger when the story does not mention whether the victim was armed, and smaller when it depicts the victim as armed.[5] When no weapon is reported, our findings imply that the use of obfuscatory language decreases the assigned responsibility and the desired level of accountability for the officer; and increases how well justified the officer was judged to be in the shooting. When a weapon is reported present, we find smaller, but still statistcially significant, effects for sentence structures that do not explicitly identify the police officer as a subject (*no subject* and *intransitive*), but statistically

[5] As expected, we also find evidence that the reported presence of a weapon matters directly, decreasing responsibility and desired accountability for the officer.

insignificant effects for the use of obfuscatory language overall, as the effect of *passive voice* is weaker in this case. Overall, the experimental results indicate that there are strong effects of narrative structures on participants' judgments about the police officer's actions in the specific incident described in the news story. We also find that obfuscating the direct role of the police in the killing reduces donations for reform by a modest amount.

With the experimental results in hand, we close the paper by turning back to the news data to examine whether the media is more likely to use obfuscatory language in high leverage circumstances when language might matter more for perceptions. Media obfuscation is indeed more prevalent for police killings in which the victim was unarmed–that is, precisely when our experimental results imply such language works to soften judgments about the moral responsability of the police officer for the killing. We also find more obfuscation in police killings for which body-worn camera footage is available, again suggesting such language is used in cases where viewers are potentially more likely to form harsher judgments against the police.

Our paper relates to several strands of prior literature. First, our work contributes to a growing literature that employs natural language processing and computational linguistics to analyze financial reports, newspaper articles, press releases, opinions, social media comments, or congressional transcripts (See Gentzkow *et al.* 2019). These approaches have been fruitful to understand how the tone of a text or speech impacts political outcomes (Grimmer, 2010; Gentzkow & Shapiro, 2010), firm performance and firm exposure to political, social, and climate risks (Baker *et al.*, 2016; Hassan *et al.*, 2019; Engle *et al.*, 2020; Giglio *et al.*, 2021). These studies used natural language processing methods to capture political slant, company executives' views, and market participants' sentiments. Our paper adds to this literature by exploring a new aspect of language - semantics - in a new context: the use of obfuscatory language in news media.

Our paper also contributes to the literature documenting the impact of media on a variety of economically and socially relevant outcomes.[6] The content and presentation of news has been shown to affect health choices (Bursztyn *et al.*, 2022b), financial markets (Baker *et al.*, 2016; Engle *et al.*, 2020), and attitudes towards immigration (Gentzkow & Shapiro, 2004; DellaVigna *et al.*, 2014; Djourelova, 2021); and that prospective media coverage influences politicians' actions (Durante & Zhuravskaya, 2018). We add to this literature by

[6]For brevity, we focus on the effects of news coverage, but recognize that there is a long literature on other dimensions of media, such as entertainment TV (Kearney & Levine, 2015), movies (Dahl & DellaVigna, 2009), or educational programming (Gentzkow & Shapiro, 2008). See DellaVigna & La Ferrara (2015) for a fuller review of the literature.

documenting differences in the media's semantic choices in coverage of police killings and analyzing how these choices affect perceptions about these incidents. Closer to our specific research question, past research shows that news also influences people's voting behaviors (DellaVigna & Kaplan, 2007; Chiang & Knight, 2011; Cagé, 2020; Couttenier *et al.*, 2021). Our online experiment shows that information and narrative structures influence people's perception of events, which in turn could affect political stances and support for policies (Bursztyn *et al.*, 2022a; Alesina *et al.*, 2021; Andre *et al.*, 2021). Several papers also relate news coverage or political speeches to perceptions of crime and police or jury behaviors (Mastrorocco & Minale, 2018; Mastrorocco & Ornaghi, 2020; Philippe & Ouss, 2018; Grosjean *et al.*, 2022).

Past research outside of economics has also investigated how police departments' activities are covered in the news, mainly through the coverage of crime. For example, in early studies, Gilliam & Iyengar (2000) and Gilliam Jr *et al.* (1996) show overreporting of crime when the suspect is Black, while later studies do not find this to be the case (Dixon & Williams, 2015). Duxbury *et al.* (2018) show that media is more likely to emphasize mental illness if the perpetrator of a mass shooting is White instead of Black or Hispanic. Grunwald *et al.* (2022) find that police agencies' Facebook posts overrepresent Black suspects relative to local arrest rates. A few papers have also explored variation in media coverage of crime depending on the gender of the perpetrator or victim (Frazer & Miller, 2009; Henley *et al.*, 1995; Yasmin, 2021), and documenting the high use of passive structures when talking about sexual assault (Bohner, 2001; Lussos & Fernandez, 2018).

Lastly, our paper builds on findings in cognitive science and linguistics. Closest to our work, past research has shown that language choices affect perceptions of the moral responsibility of perpetrators (De Freitas *et al.*, 2017) and victims (Henley *et al.*, 1995; Niemi & Young, 2016; Northcutt Bohmert *et al.*, 2019). Our experimental results provide new evidence along these lines, demonstrating that the use of obfuscatory language decreases the assignment of moral responsibility and the desired level of accountability for police officers who kill civilians.

2 Linguistics Framework

In this section, we describe the linguistics framework that forms the basis for our empirical analysis. Our primary goal is to identify particular sentence structures that work to either clarify or obfuscate the actor and actions described in a news story, thereby affecting the viewer's (reader's) perception of what happened and who was responsible. Our framework draws heavily on Chapter 8 of Toolan (2013), which includes a detailed characterization of how different narrative structures affect perceptions of causal relationships and the assignment of causal agency.[7]

In psycholinguistics a "causative construction" refers to the way in which language is used to depict causation from one subject (causal agent) to another (causal patient). The point of comparison throughout our paper is the use of an active sentence structure that clearly identifies the action, the causal agent, and the causal patient – i.e., sentences of the form: "A police officer killed a man." Following Toolan (2013), we focus on four key dimensions of sentence structure that can be used to obscure or obfuscate the action or roles of the causal agent and patient: (i) the use of passive versus active voice, (ii) nominalization - i.e., turning an action verb into a noun, (iii) the failure to identify a causal agent, and (iv) the use of intransitive versus transitive verbs. We present each of these sentence structures in turn. Table 1 provides simple examples (columns 1 and 3) and examples drawn directly from our news broadcast data (columns 2 and 4), to help illustrate these sentence structures for police killings (columns 1 and 2) and civilian killings (columns 3 and 4).

Passive versus Active Voice. The sentence "A police officer killed a man" uses the active voice and a transitive verb, identifies the causal agent as subject, and does not use nominalization. A first way to diminish the responsibility of the causal agent for the action is to instead use the passive voice: "A man was killed by a police officer." With the active voice, the subject acts upon its object through the act described by the verb. This sentence structure is considered to be strong, direct, and clear in tone. It also places the causal agent at the center of the sentence. In contrast, the passive voice relegates the causal agent (police officer) to the background of the sentence, lowering the salience to the reader.

Although the function of the passive is to background the agent, not necessarily to obfuscate, there is evidence that its use changes the perception of the reader or viewer. As noted by Chestnut & Markman (2018), "[S]tating 'The woman was abused by the man' rather than 'The man abused the woman' causes people to be more accepting of violence

[7] Toolan (2013) also discusses other forms of narrative that can modify perceptions of causal relations, such as direct commentary and evaluation or editorial choices on how to name things (for example, choosing between the terms "rioter" or "demonstrator"). While also of potential interest regarding the media coverage of police killings, characterizing these aspects of media language is beyond the current scope of our research, because it is difficult to identify such strategies at scale and because it is more challenging to compare their usage across different kinds of news stories, as these approaches are likely to be domain-specific.

against women, because passive voice distances perpetrators from their crimes and consequently makes the crimes seem less severe (Henley *et al.*, 1995)." Recent work in linguistics and cognitive science suggests that the passive increases 'psychological distance' with respect to the narrated event by making it seem more distant in time and space, as well as more hypothetical (Chan & Maglio, 2020). Furthermore, as we will see below, the passive voice also makes it easier to omit a subject altogether, underemphasizing the role of the causal agent to an even greater extent.[8]

Nominalization. Nominalization is the process of transforming an adjective or verb into a noun. It is a key linguistic resource in everyday language, as it allows one to refer to an event without fully narrating it. In news reporting, it helps shorten stories, but can also be a tool to obfuscate agency, since it puts elements of a story into the background and leaves some aspects of the narrative ambiguous. In the context of police killings, a common form of nominalization is the use of the term "officer-involved shooting" as a replacement for a more direct sentence structure like "a police officer killed a civilian". There are two things to note in this case. First, even though the participation of an officer is noted with this form of nominalization, the officer's causal role in the shooting is left ambiguous. Second, although the police officer might have killed someone (as in our data), the chosen verb for the nominalization is not 'kill', but 'shoot'.[9] Thus, it also leaves ambiguous the fact that someone was killed. In the context of civilian killings, phrases like "intimate partner killing" or "gang related shooting" can also be used.

In other instances the noun 'death' is used instead of an active verb like 'kill'. This form of nominalization tends to focus the reader's attention on the final outcome (death) rather than the action of the agent that caused the outcome. In this specific case, the word death is also derived from the intransitive verb 'to die'. As we discuss below, the use of the intrasitive works, in addition, to not only remove a causal agent from the sentence, but also to remove the need for a cause at all.

Failure to Identify the Causal Agent. A third way to diminish the ability of a reader to assign responsibility for an action is to remove the causal agent from the sentence altogether – e.g., "A man was killed following a police chase" instead of "A man was killed by a police officer." In this case, the person responsible for the killing is not linked directly to the action, often leaving some ambiguity about who was responsible. This is in contrast

[8]For a typological and functional overview of the passive, see Kazenin (2001).
[9]The comparable phrase 'officer-involved killing' is not commonly used by the news media.

to cases in which the causal agent is the subject of a sentence (as in the active voice) or when the passive explicitely mentions the causal agent after a causative preposition (such as 'by'), where a direct connection between the agent and the action is made.

Intransitive versus Transitive Verbs. A fourth way to obfuscate causation and responsibility is to use an intranstive verb: "A man died following an incident on the North Side." Transitive verbs like 'to kill' or 'to shoot' imply that somebody caused another individual to die, even when the causal agent is not clearly identified as the subject of the sentence. Wolff (2003) and Pinker (2007) categorize transitive verbs like 'kill' and 'shoot' as causative verbs, because they implicitly relay the idea that the causal agent in the sentence caused an event directly, intentionally, and without an intervening actor. By contrast, intransitive verbs like 'to die' do not require a causal agent or even a cause at all. Instead intransitive verbs only need mention the causal patient, here the deceased person. Thus, the use of intransitive verbs not only obfuscates who was responsible for an event, but implicitely rejects that there is a causal agent, directly increasing the ambiguity about what happened in the first place.[10]

Putting It All Together. For simplicity, we use the labels *Passive*, *Nominalization*, *No Agent*, and *Intransitive* to refer to these four forms of obfuscatory sentence structure throughout the rest of the paper. Importantly, these sentence structures are not mutually exclusive and are often combined in practice. We use the term *Any Obfuscation* to refer to the use of any of these structures and define a sentence as having *No Explicit Agent* if *No Agent* or *Intransitive* applies. Finally, as the progression in Table 1 suggests, we use the order in which we introduced these sentence structures - *Active* $>>$ *Passive* $>>$ *Nominalization* $>>$ *No Agent* $>>$ *Intransitive* - to define a hierarchy of causal clarity or, in reverse, a hierarchy of obfuscation.

[10] Note that you can not say 'The police officer *died* the victim' or 'The victim was *died* by the police'. Pinker (2007) refers to this as the intransitive "resisting" a causative.

Table 1: Obfuscatory Sentence Structures - Simple Examples

	Police killing		Civilian killing	
Semantic structure	Simple example	News broadcast	Simple example	News broadcast
Active	A police officer killed a person	Investigators believe the police officer shot and killed the man just before nine o'clock	A suspect killed a person	They believe [name of suspect] stood over the boys as they slept on the couch and shot them
Passive	A person was killed by a police officer	New developments, a California man under arrest tonight accused of making a prank call that led to a victim being shot to death by the police	A person was killed by a suspect	The Goodhue County attorney says that the man was shot in the chest by the suspect early yesterday morning.
Nominalization	A person was killed in an officer-involved shooting	An officer-involved shooting late Thursday night claimed the life of a Monroe man	A person was killed in a domestic-violence shooting	A 30-year old man was shot and killed on Tuesday in a gang-related shooting.
No Agent	A person was killed	Officials say the 45-year old man was shot after he refused to drop a knife	A person was killed	Several shots were fired at the doorway into the apartment with several adults, a toddler and an infant inside.
Intransitive	A person died	The man has died after a shoot out with police officers in Saint Louis	A person died	[Name of victim] died at the scene of that shooting.

Prior research in linguistics and moral philosophy in particular has demonstrated that variation in these sentence structures influences how readers interpret an event. For example, in early work, Trew (1979) argues that news writing uses narrative structures that reflect dominant social beliefs. Wolff (2003) and Pinker (2007), among other scholars, emphasize how perceptions of causation can be influenced by choices in sentence construction. Closely related to our work, De Freitas *et al.* (2017) shows that there is a close relation between choices of causative verbs and the subsequent moral judgment of viewers/readers.

3 Data, Sample Construction, and Language Processing

3.1 Primary Data Sources

We draw data from several sources: a comprehensive dataset on the universe of police killings in the US between 2013 and 2019; a database that includes the (almost) universe of gun-related killings in the US from 2014 to 2018, and the closed captions (text transcriptions) of all televised news broadcasts on both local and national stations from 2013 to 2019[11].

Police killings. There is no official government record of police killings in the United States. As a result, in recent years, journalists, activists, and researchers have undertaken independent efforts to build a comprehensive database of all such killings. For our analysis, we use data from *Mapping Police Violence* (MPV). The MPV research collaborative identifies and documents all police killings in the US, starting in 2013. The incidents are identified from other crowdsourced databases on police killings in the United States, including FatalEncounters.org. MPV processes each potential case, and improves the quality and completeness of the data by examining available information about the case in the traditional media, social media, and obituaries. Conner *et al.* (2019) finds that the MPV had a coverage of 98.3% of all police killings in 2015. The MPV dataset includes information on the name of the victim, the police department responsible for the killing, the race of the victim, and the address and zip code where the incident took place.

Civilian gun homicides. There is also no official government database of civilian homicides in the United States. To build a database of civilian homicides comparable to the

[11] As mentioned below, audits of the our police killings data shows a coverage of 98.3% of all cases.

police killings identified in MPV, we draw on data from the *Gun Violence Archive* (GVA). This database is collected by a non-profit organization aiming at registering all known shootings in the country. Incidents in the GVA are collected daily from over 7,500 law enforcement, media, government, and commercial sources. Each incident is verified by both an initial researcher and a secondary validation process. Just as the MPV, the GVA includes the name of the victim in the vast majority of cases. The GVA also includes information on the suspect, when available. Since the GVA does not include information on the race of the victim or suspect, we impute the posterior probability that each subject belongs to a particular racial or ethnic group with the information on the name and location where the incident took place (Imai & Khanna, 2016; Khanna *et al.*, 2017).[12] Finally, to isolate only civilian homicides in the GVA, we drop all suicides and accidental deaths as well as deaths due to a police shooting.

Television news broadcasts. Our media dataset contains the universe of closed caption text across all television news programs in the United States. The data was provided by *News Exposure* (NE), a data vendor that monitors and collect transcripts from over 800 distinct TV stations across the 210 media markets in the U.S. Both local and national stations are included in the database.[13] Altogether, over 2 million station-days of news transcripts were available for our analysis. As we describe below, we searched these comprehensive television news transcipts for stories about the police killings and civilian homicides recorded in the MPV and GVA, respectively. In addition to the text of an associated news story, we obtained information on the station, network affiliation, media market, date and time of the broadcast, run time, publicity value, ad value and ratings estimate.

We complement the previous data sources with information on the demographics of the tract and media market in which the killings took place from the American Community Survey and Census. We also merge Designated Media Area (DMA) demographics and electoral results from Martin & McCrain (2019).

[12]The algorithm uses the probability of being part of a racial or ethnic group based on the name and Census tract, with a 50 percent probability threshold as in Moreno-Medina (2021) and Humphries *et al.* (2019). This posterior probability is estimated using the package WRU in R.

[13]These data are also used in Moreno-Medina (2021).

3.2 Sample Construction

To measure the use of obfuscatory language structures in media coverage of police killings, we need a meaningful benchmark, as absolute levels of language usage are difficult to interpret. To this end, our primary analysis compares media coverage of police killings to civilian homicides. We make several sample restrictions to ensure this comparison is as meaningful as possible.

First, since the GVA database only includes information on gun deaths (not other forms of homicide), we limit our sample to police killings caused by gunshot, which represent more than 90% of all police killings.[14]

Second, we make a set of sample restrictions designed to isolate circumstances in which the media *could have used similar language* to describe police killing and civilian homicides. A key issue is that while it is possible to identify a police officer as the causal agent in most police killings, it may not be possible for the media to identify a causal agent in a civilian homicide when a suspect/perpetrator has not yet been identified. In the absence of such information, it is natural for the media to instead focus the narrative on the victim - e.g., "a 40-year old man was shot last night". To make the police and civilian killings as comparable as possible, therefore, we limit our sample of civilian homicides to those in the GVA database in which the name of the suspect is known.[15]

For our baseline analysis, we further limit our sample to news stories in which the suspect's name appears. To avoid concerns that this sample restriction biases our analysis towards finding greater use of active sentence structure for civilian homicides, we consider a number of alternative specifications to ensure robustness, including dropping sentences that include the suspect's name and focusing on the first sentence in the story.

Appendix table D.1 shows how case composition changes with our sample construction choices. For police killings, our sample is very similar on all observables. For civilian killings, the requirement that the suspect is known increases the share of domestic violence and murder-suicide cases in the data. In turn, our sample has more women, and the victims are a little older than the average shooting victim in the United States. The racial and geographic composition of victims in our analysis sample are the same as the full sample.

[14]In particular, of the 7,663 police killings documented in MPV, 7,299 are caused by gunshot and our final sample consists of the 7,293 of these that could be geolocated.

[15]Starting from an original sample of 49,277 gun deaths, we drop all suicides and accidental deaths, drop deaths due to a police shooting, and restrict the sample to those civilian homicides in which the name of the suspect is available in the GVA. These sample restrictions yield a dataset of 19,325 civilian gun homicides and our final sample consists of the 17,939 of these that could be geolocated.

Table 2: Descriptive Statistics by Individual

	All	Police Killings	Civilian Killings
	(1)	(2)	(3)
Panel A: Subject Level			
Victim Chars:			
Age	35.13	36.79	33.90
Male	0.81	0.95	0.72
Black	0.24	0.18	0.28
Hispanic	0.12	0.16	0.09
White	0.52	0.58	0.47
Other/Unknown	0.12	0.07	0.15
Incident Chars:			
Body Camera	0.11	0.11	.
Victim Not Fleeing	0.66	0.66	.
Share Vote Rep. DMA	0.51	0.49	0.52
Observations	13,702	5,759	7,943
Panel B: Sentence Level			
Obfuscation Dims.:			
Passive	0.20	0.22	0.17
Nominalization	0.04	0.04	0.03
No Agent	0.16	0.16	0.15
Intransitive	0.11	0.12	0.11
Any Obfuscation	0.34	0.36	0.29
No Explicit Agent	0.15	0.15	0.14
Observations	466,636	320,042	146,594

Source: MPV; News Exposure. *Notes:* This table presents the mean of different variables of cases in our sample. The coverage information is at the sentence level. Our sample includes sentences where both a victim and an alleged perpetrator were identified. Data sources: GVA, MPV and News Exposure.

To match police killings and civilian homicides to news coverage, we use a machine learning-based procedure that follows three sets of requirements intended to ensure a high quality match, as in Moreno-Medina (2021). First, we subset the NE data to text transcriptions that include words related to a killing/homicide such as "shot", "shoot", or "killed", which sharply increases the probability that a news story is about crime.[16] We keep stories

[16]For civilian homicides, we use all forms of the following keywords: shot, gunshot, kill, and homicide. For the police shootings, we search for all forms of the keywords: shot, gunshot, kill, homicide, police, and officer.

with a score above a certain threshold and manually check the accuracy of this threshold, finding that 99% of all the identified stories are indeed covering a crime or police incident.[17]

Second, we require a story to contain either the name of the victim or the address (block and street) in which the event happened. Third, we consider only stories that aired within 7 days of the victim's death. The goal of this last restriction is to limit misclassification of stories (especially for victims with common names) by essentially requiring a match on both name/address and date.Finally, for our analyses that are at the sentence level, we focus on sentences in which: i) there are references either to the victim or the suspect and; ii) the sentence is informing on the killing. See Section 3.3 for further details. Our final sample includes 192,944 stories and 466,636 sentences linked to 5,759 police gun killings and 7,943 civilian gun homicides for which we were able to find at least one broadcast news story.[18]

The first two panels of Table 2 present descriptive statistics for the police killings and civilian homicides samples. Compared to victims of civilian gun homicides, victims of police shootings are much more likely to be male (95% compared to 72%), are slightly older (38 years old compared to 34 years old), more likely to be Hispanic (16% compared to 9%) and White (58% compared to 47%), and less likely to be Black (18% compared to 28%). We control for these demographic variables in all of our empirical analyses. Appendix Figure C.1 presents trends over time in the number of police killings and civilian homicides included in our final sample. Overall, our sample includes more cases of civilian homicides, but more stories about police killings.

3.3 Language Processing

In this section we explain how we process the text data once we have our primary sample of news stories linked to police killings and civilian homicides. Appendix figure A presents a flowchart of our process. We apply three language processing steps to construct the measures of obfuscation. Our implementation of these steps uses state-of-the-art Natural Language Processing models, based on *BERT* (Bidirectional Encoder Representations from

[17] In this way, our algorithm does a good job of ruling out unrelated stories that might uses similar words - e.g., a sports story in which the word 'shot' describes a basketball or soccer play rather than the action of a gun.

[18] Note that there are on average 56 stories per police killing, compared to 19 per civilian killing. In section 4.2, we provide evidence that our results are not driven by differences in the volume of coverage.

Transformers).[19] As will be clear below, we need a language model like *BERT* that captures contextual embeddings for words. The three steps are:

1. Co-reference resolution: identify all words that reference the same individual
2. Identify who did what to whom in sentences about the shooting
3. Encode our measures of obfuscation

Additional details can be found in Appendices A.1 to A.4.

Co-reference resolution and story delimitation. First, we identify all the words that are used to refer to the same individual (victim or suspect), including pronouns. We adopt the method proposed by Lee *et al.* (2018), and implemented by Gardner *et al.* (2017), which uses a BERT-like neural network called SpanBERT (See Appendix A.2 for more details). The model takes text as an input and outputs a list of clusters of tokens (or words) that are considered to refer to the same individual. We define a story about the killing as the span between the first and last sentence in which the victim or suspect appears in the caption.

Semantic role labeling. Second, we need to identify for each sentence who (agent) does what (verb) to whom (patient). This task is known in the NLP literature as Semantic Role Labeling (Appendix A.3 provides more details). We implement another BERT-type model, this time the one proposed by Shi & Lin (2019). Given that we want to identify how the killing is being covered, we focus on sentences that include verbs informing on the killing ('kill', 'shoot', etc). For those sentences the algorithm produces an analysis for each verb, detailing who is executing the action, and who is being acted upon. For our purposes, we want to identify who is executing the action of killing or shooting (agent) and who is being killed or shot (patient). This same output allows us to see if an individual is the subject of the intransitive 'die' as well. We classify each verb into the following categories:

- Transitive: verbs that either start with any of the texts in the following list: 'kill', 'shot', 'gun', 'murder', 'shoot', 'hit', 'fire', 'open', 'strike'; are passive conjugations of

[19] *BERT* (Bidirectional Encoder Representations from Transformers) is a neural network model of language that has proven to be incredibly successful among a host of tasks in natural language processing. There are several technical features in BERT, but perhaps the most important is that it trains the model not only using previous words in a text, but also future ones. The standard model allows up to 512 words (or tokens) in a text. The network has 7 layers, and it works with a type of word embedder model that captures the context in which the word is being used. Since 2019, *Google Search* has been applying BERT models for English language search queries within the US.

the above verbs;[20] or past participle verbs 'declared', 'found' or 'pronounced' followed by the past participle of shoot or kill - shot or killed.

- Intransitive: verb 'die', or auxiliary verb followed by the past participle of die - died
- Irrelevant: all others

We focus on sentences in which the identified patient for these verbs is the victim in our data.

Obfuscation classification. Lastly, we create our measures of obfuscation as described in Section 2; that is, we define whether each sentence presents any of the following structures: active, passive, nominalization, no agent, intransitive. Appendix A.4 presents the exact phrases used.

4 Obfuscation in news stories on police killings

We now present the results of our analysis of news broadcasts examining whether the media is more likely to use obfuscatory sentence structures in stories about police killings compared to a control sample of civilian homicides. In the first subsection, we present two analyses using *all* and *first* sentences within identified news stories that reference the killing/death. We focus, in particular, on the first sentence or 'lead', because it is likely the most salient to viewers. In the next subsection, we present results of a number of additional specifications designed to examine whether the main results are robust to alternative ways of designing the study, primarily related to sample selection.

4.1 Main results

To get an initial sense of whether the news media is more likely to use obfuscatory sentence structures in stories about police killings, Panel B in Table 2 presents summary statistics on the prevalence of different structures broken down by whether a police officer or civilian was responsible for the killing. Overall, in the raw data, the use of any obfuscatory sentence structure was 24 percent more likely when a police officer was responsible for the killing (36% versus 29%). This aggregate result reflects the greater use of all four obfuscatory

[20] That is, the verb is 'to be' or other auxiliatory verb, but is followed by the past participle of the above verbs [for example, 'the man was killed'].

sentence structures in stories about police killings, especially the passive voice (22% versus 17%). Appendix Figure C.2 plots the fraction of sentences with obfuscation over time, revealing that the increased prevalence of obfuscation in stories about police killings is stable over the study period.

Table 3: Obfuscation in news of police killings: sentence-level analyses

Outcomes: Dimension of Obfuscation	Mean Civ. Shoot.	Coefficient on Police Killing Indicator (1)	(2)	(3)	(4)	(5)
Panel A: Aggregate Dimensions:						
Any	0.2935	0.062***	0.070***	0.073***	0.074***	0.075***
		(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
No Explicit Agent	0.1354	0.017***	0.020***	0.025***	0.026***	0.032***
		(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
Panel B: Individual Dimensions:						
Intransitive	0.1075	0.011**	0.013***	0.017***	0.018***	0.023***
		(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
No Agent	0.1497	0.011**	0.016***	0.016***	0.015***	0.010***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Nominalization	0.0282	0.010***	0.011***	0.012***	0.012***	0.013***
		(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Passive	0.1692	0.047***	0.052***	0.051***	0.051***	0.046***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Story Controls			X	X	X	X
DMA FE				X	X	X
Station FE					X	
Month-Year FE						X
Observations		466,636	466,636	466,636	466,636	466,636

Notes: Standard errors clustered by subject (*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$). This table presents differential obfuscation for stories about police killings relative to civilian killings from Equation 1. Our analyses are at the sentence level. We vary what controls are included across columns, including a time trend. Each row presents a separate regression coefficient on a dummy equal to 1 if the story is about a police killing, 0 if it involves a civilian killing, for different obfuscation measures described in the first column. Our sample includes incidents and news stories where a suspect was identified for civilian killings. All sentences include some mention of either the victim or suspect. We define 'obfuscation' as having passive forms, no agent, intransitive verbs, or nominalization. We define 'No Explicit Agent' as having no agent, intransitive verbs, or nominalization. See Section 2 for more details. Source: News Exposure.

To control for other potential differences in stories about police and civilian killings - e.g., the age, sex, and race of the decedent - we estimate regressions of the form:

$$Obfuscation_{eitsd} = \beta_1 Police_i + \beta_2 X_{eitsd} + \epsilon_{eitsd} \quad (1)$$

where e indexes a sentence about incident i, at time t on station s in media market d. $Police_i$ is a dummy equal to 1 if the news story is about a police killing, X_{eitsd} are controls for characteristics of incident, date, television station, and media market. All of our analyses, except the one without any controls, include a linear time trend.

For our first analysis, we treat each sentence in the story that references a killing/death as an observation and cluster the standard errors at the individual subject level. Table 3 presents results for five specifications. The specification shown in Column 1 includes no controls. Column 2 adds story-level controls (age, sex, and race of the victim), while Columns 3 and 4 successively add media market and television station fixed effects. The specification shown in Column 4 is our preferred specification because it effectively compares the coverage of police versus civilian killings by the same television station in the same media market. Column 5 replaces the linear time trend for the specification in Column 3 with month by year fixed effects, primarily to check whether there are any non-linear time effects for which the linear time trend is not controlling adequately.

For each of these five specifications, we report results for six dependent variables in the rows of the table. The final four rows report the results for the four distinct obfuscatory sentence structures described above, while the first two rows report results for dependent variables that aggregate these outcomes. The second row reports results for the aggregate category, *No Explicit Agent*, which combines *No Agent* and *Intransitive*, while the first row aggregates all four categories to report the use of *Any Obfuscation* in the sentence. We find a consistent pattern of results across all of the specifications shown in Table 3: sentences in stories about police killings are around 25 percent (7 percentage points) more likely to use some form of obfuscation compared to stories about civilian killings and there is an increased propensity to use each distinct form of obfuscation (rows 3-6).

In our preferred specification (column 4), *Passive* and *No Explicit Agent* sentence structures are 30 percent (5.1 percentage points) and 19 percent (2.6 percentage points) more likely to be employed in stories about police killings, respectively. *Nominalization* is generally the least common form of obfuscation to be used by the media, but it too is much more prevalent (43 percent, 1.2 percentage points) in stories about police killings. Overall, our results indicate that news coverage of police killings is significantly more likely to use narrative structures that obscure the police officer's responsibility for the killing relative to civilian homicides.

Table 4: Obfuscation in news of police killings: first sentence in a news story

Outcomes: Dimension of Obfuscation	Mean Civ. Shoot.	Police Killing (1)	(2)	(3)	(4)	(5)
Panel A: Aggregate Dimensions:						
Any	0.2814	0.105***	0.118***	0.120***	0.120***	0.119***
		(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
No Explicit Agent	0.1266	0.046***	0.053***	0.059***	0.060***	0.066***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Panel B: Individual Dimensions:						
Intransitive	0.0748	0.035***	0.041***	0.046***	0.046***	0.051***
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
No Agent	0.1405	0.022***	0.027***	0.025***	0.024***	0.015**
		(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
Nominalization	0.0525	0.019***	0.020***	0.021***	0.022***	0.025***
		(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Passive	0.1663	0.063***	0.069***	0.066***	0.065***	0.058***
		(0.007)	(0.008)	(0.008)	(0.007)	(0.007)
Story Controls			X	X	X	X
DMA FE				X	X	X
Station FE					X	
Month-Year FE						X
Observations		182,142	182,142	182,142	182,142	182,142

Notes: Standard errors clustered by subject (*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$). This table presents differential obfuscation for stories about police killings, relative to stories about civilian killings, from estimating Equation 1. Our analyses just for the first sentence of each news story. We vary what controls are included across columns, including a time trend. Each row presents a separate regression coefficient on a dummy equal to 1 if the story is about a police killing rather than a civilian killing, for different measures of obfuscation, which are described in the first column. Our sample includes incidents and news stories where a suspect was identified for civilian killings. All sentences include some mention of either the victim or suspect. We define as having passive forms, no agent, intransitive verbs, or nominalization. We define as having no agent, intransitive verbs, or nominalization. See Section 2 for more details. Source: News Exposure.

Table 4 presents an analogous set of results to those reported in Table 3 for specifications that include only the first sentence of the story. News organizations generally present what they consider to be the essential or attention-getting facts about a story in the first sentence, which is generally expected to be especially salient to viewers (AP, 2020). As a result, we expect any obfuscation in the 'lead' to have an outsized effect on how viewers understand and respond to the incident.

The results of the analysis of first sentences are qualitatively similar and quantitatively greater than those presented in Table 3. In this case, obfuscation is about 40 percent (12 percentage points) more likely for coverage of police killings versus civilian homicides. For first sentences, *Passive*, *No Explicit Agent*, and *Nominalization* are each about 40-50 percent more likely to be used in stories about police killings. Overall, the results presented in Table 4 suggest that the media's use of obfuscation in the coverage of police killings is especially likely in the first sentence - i.e., most salient part of the story.

4.2 Robustness checks

Table 5 reports the results of a number of specifications designed to examine the robustness of our main findings to alternative ways of designing the study. For comparison, Panel A repeats the estimates from our preferred specifications (column 4) in Tables 3 and 4, while Panels B-D report analogous results for three alternative models.

In constructing the control sample of civilian homicides for our analysis, our goal was to identify situations in which the media faced a similar choice of language for both police and civilian killings. As a key sample selection criteria, we require the suspect's name to appear in the story. This choice was made to ensure that an agent (police officer, civilian suspect) could possibly have been used in the story - i.e., to rule out cases where a death occurred but nothing about a potential suspect (or even whether the incident was a homicide) was known at the time of the news report. One of our primary concerns with this sample selection criteria, however, is the possibility that requiring the suspect's name to appear in the story might bias our sample towards including more active sentence structures for civilian homicides. We were particularly concerned that the suspect's name might appear commonly as the subject of a sentence describing the murder.

To address this concern, Panel B reports the estimates for a specification that drops all sentences that include the suspect's name. The results are virtually unchanged, suggesting our concerns about the use of more active sentence structures involving the suspect's name were unfounded. Interestingly, only 1,735 sentences are dropped in this specification, which is much smaller than the total number of stories (58,033) about civilian homicides, despite our criteria that the suspect's name appears in the story. This implies that most of the time the suspect's name appears in a news story, it is not in a sentence included in our main analysis (which must directly describe the death/killing). Instead, the suspect is often named in a a stand-alone sentence - e.g., John Doe was identified as the suspect in the case. - and, thus, does not make it into our main analysis sample.

Panel C reports results for specifications which remove news stories about domestic violence. Our concern in this case is that news stories about domestic violence might be especially likely to center around the victim, resulting in the use of different sentence structures. The results reported in Panel C are largely unchanged.

Table 5: Robustness tests

	All Sentences		1st Sentence	
Dimension of Obfuscation:	Any (1)	No Explicit Agent (2)	Any (3)	No Explicit Agent (4)
	Panel A: Main results (Panels D of Table 2 and 3)			
Police Killing	0.074***	0.026***	0.120***	0.060***
	(0.008)	(0.005)	(0.009)	(0.006)
Observations	466,636	466,636	182,142	182,142
Mean Civ.Shoot.	0.2935	0.1354	0.2814	0.1266
	Panel B: Dropping sentences where suspect is named			
Police Killing	0.075***	0.027***	0.121***	0.060***
	(0.008)	(0.005)	(0.009)	(0.006)
Observations	464,901	464,901	181,489	181,489
Mean Civ.Shoot.	0.2928	0.1350	0.2803	0.1259
	Panel C: No domestic violence in civilian shooting			
Police Killing	0.069***	0.026***	0.117***	0.058***
	(0.008)	(0.005)	(0.009)	(0.006)
Observations	434,359	434,359	170,009	170,009
Mean Civ.Shoot.	0.2983	0.1369	0.2829	0.1277
	Panel D: Weighted by 1/# Sentences			
Police Killing	0.073***	0.047***	0.111***	0.057***
	(0.005)	(0.004)	(0.007)	(0.005)
Observations	466,636	466,636	182,142	182,142
Mean Civ.Shoot.	0.3237	0.1580	0.3062	0.1483
Controls	Story+DMA FE+Station FE			

Notes: Standard errors clustered by subject (*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$). This table presents robustness tests for our main results, for all sentences (columns 1 and 2) and for the first sentence (columns 3 and 4). Panel A presents our preferred specification from tables 2 and 3 (column 4). In Panel B, we drop sentences in the civilian killing sample where the suspect is named. In Panel C, we drop stories about domestic violence incidents. In Panel D, we reweight sentences by 1/number of sentences in a particular story. We define 'Any (obfuscation)' as having passive forms, no agent, intransitive verbs, or nominalization. We define 'No Explicit Agent' as having no agent or intransitive verbs. See Section 2 for more details. Source: News Exposure.

Finally, we provide two tests to see whether the difference in language is due to the difference in volume of coverage per story. As specified earlier, there are on average 56 stories per police killing, against 19 stories per civilian killing. One may be concerned that a small number of viral incidents that receive a great deal of media attention are driving our main results. First, Panel D reweights each sentence by the inverse of the total number of sentences per victim - thereby, giving equal weight to all victims. The findings are again remarkably similar to our main results, implying that the increased use of obfuscatory sentence structures for police killings is not limited to high profile cases. Second, the first analysis in Appendix Figure C.3b breaks out our sample by whether an incident led to "viral" media coverage or not with viral coverage being defined as having more than 100 news segments. For non-viral incidents, there are on average 16 stories per civilian killing, compared to 27 per police killing (these numbers are 447 and 620 for viral stories). Results are very similar in these two kinds of stories, suggesting that volume of stories per incident is not driving differential obfuscation. [21]

5 The Effects of Obfuscatory Language

The analysis presented in Section 4 reveals the systematic use of more obfuscatory language in broadcast news coverage of police killings relative to civilian homicides. But does this matter in practice? It could be argued that the sentences `'a police officer shot and killed a man'` and `'a man died in an officer-involved shooting'` contain the same information. Does the difference in language really affect the way viewer or readers understand and respond to a news story? To test this, we conduct an online experiment to measure how the narrative structure used to describe a police killing affects a viewer's assessment of the officer's moral responsibility, demand for accountability for the officer, broader support for police reform, and subsequent re-telling of the story.

5.1 Experimental Design

We conducted an online experiment with 2,402 participants in March 2022 using Prolific. Our hypotheses and research and analysis design were registered on the AEA registry (AEARCTR-0009052). Participants were required to reside in the United States and to be

[21] Appendix Table D.2 provides further robustness tests. In Panel A, we present results evaluated at the story level, instead of the sentence level. In Panel B, we first limit our sample to years in which we have data from both GPA and MPV; on the second check, we limit our sample to cases that appear on the news for at least two separate calendar days. The results are again unchanged in these additional robustness checks.

Table 6: Online experiment: First sentence of the news story for each narrative treatment arm

Narrative treatment arm	Headline
Active	A police officer killed a 52-year-old man on Friday night.
Passive	A 52-year-old man was killed by a police officer on Friday night.
No Agent + Nominalization	A 52-year-old man was killed in an officer-involved shooting on Friday night.
Intransitive + Nominalization	A 52-year-old man died in an officer-involved shooting on Friday night.

adults fluent in English. [22]

We presented participants with a story (a headline sentence plus four sentences providing further detail) about a police killing. Participants all read about the same incident, but were randomly assigned to variations in how it was described, using a 4 X 2 design. The first level of randomization was for narrative structure. Participants were randomized to one of four structures: (1) *Active*, (2) *Passive*, (3) *No Agent + Nominalization*, and (4) *Intransitive + Nominalization*.[23] Table 6 provides the headline sentence used in each of these sentence structure treatment arms; the full prompts can be found in Appendix B. Following the definitions used earlier in the paper, we define *Any Obfuscation* as having grammatical structures (2), (3) or (4) and *No Explicit Agent* as having narrative structure (3) or (4). The second level of randomization determined whether a clause stating that the man killed "was reportedly armed" was included in the story or not.

We are interested in understanding how narrative structure influences three broad sets of outcomes. First, does it influence how someone understands and judges the specific event being described? We study this by asking participants questions on perceptions of the officer's moral responsibility for the civilian's death and demand for penalties for the officer. Second, does it alter someone's broader understanding of police harms and their support

[22]The survey took on average six minutes to complete, and participants were paid $1.70 to participate. Appendix Table D.3 presents balance tables that confirm that randomization worked properly.

[23]For expositional ease, we refer to the latter categories as *No Agent* and *Intransitive* throughout the remainder of this section, but it is important to note that both include the language "officer-involved shooting" in the story.

for police reform? To get at this, we asked respondents how they would like to split a potential $100 donation between two organizations: one supporting officer well-being and the other supporting police reform. We use donations to the latter as our primary measure of support for reform. Finally, does narrative structure affect the way that respondents recall and re-tell the story? We study this by measuring both information content (i.e., whether they report that a police officer was responsible for the killing) and narrative structure used by participants in their recounting of the story at the end of the experiment. Our exact questions can be found in Appendix B.2.

We hypothesized that obfuscation matters – that is, respondents would be less likely to find the officer to be morally responsible, to demand accountability, and to support police reform when the news story presented some obfuscation, relative to it being told with an active voice. In addition, we hypothesized that the degree of obfuscation, outlined in section 2, is also important: the greater the obfuscation, the less likely people are to assign responsibility, demand accountability, or support reform. Lastly, we hypothesized that these effects would be strongest if we did not specify that the victim was armed.

5.2 Experimental Results

Judgments about Officer's Actions. Table 7 presents the main results of the experiment. Our primary outcomes of interest are whether the officer is morally responsible for the killing (columns 1-3), support for penalties for the officer from their police department (columns 4-6) and support for broader legal penalties for the officer (columns 7-9). For each outcome, the three columns present the effect of: (i) *Any Obfuscation*, (ii) *No Explicit Agent*, (iii) *Passive*, *No Agent*, and *Intransitive*, using Active as the reference group for all three columns.

In line with our main hypothesis, how the story is told matters for perceptions of what happened. In particular, removing the mention of an explicit agent reduces perceptions of an officer's responsibility, as well as demand for penalties. The *No Explicit Agent* treatment decreases the share responding that the officer was morally responsible by 13% (9 percentage points, $P < 0.001$); decreases the stated preference for departmental penalty by 7% ($P = 0.001$), and for legal penalties by 8% ($P < 0.001$). In contrast, we find no significant responses across the three measurements between the *Passive* and *Active* arms. As a result, the statistical significance of Any Obfuscation only rises to the 10 percent level for the outcomes related to demanding penalties for the officer.

We also hypothesized that specifying whether the victim was armed or not would sub-

Table 7: Online experiment: narrative structure and judgment of the event

	Moral Reponsibility			Department Penalty			Legal Penalty		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Obfuscation	-0.06***			-0.19*			-0.18*		
	(0.02)			(0.10)			(0.10)		
No Explicit Agent		-0.09***			-0.28***			-0.29***	
		(0.02)			(0.10)			(0.10)	
Passive		-0.02	-0.02		-0.00	-0.00		0.03	0.03
		(0.03)	(0.03)		(0.12)	(0.12)		(0.12)	(0.12)
No Agent + Nominalization			-0.06**			-0.28**			-0.31***
			(0.03)			(0.12)			(0.12)
Intrasitive + Nominalization			-0.11***			-0.27**			-0.26**
			(0.03)			(0.12)			(0.12)
Mean Dep. Var.	0.72	0.72	0.72	3.93	3.93	3.93	3.73	3.73	3.73
SD Dep. Var.	0.45	0.45	0.45	2.15	2.15	2.15	2.16	2.16	2.16
N	2402	2402	2402	2402	2402	2402	2402	2402	2402

Notes: Obfuscation is equal to 1 if the sentence structure is passive, no agent, or intransitive. No explicit agent is equal to 1 if the sentence structure is no agent or intransitive. The outcome in columns 1-3 is a dummy equal to 1 if the respondents think that the police officer is morally responsible for the victim's death. The outcomes in the remaining columns are support, on a scale from 1 to 7, for department penalties (columns 4-6) and legal penalties (columns 7-9), respectively. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

stantially influence respondents' perceptions about whether the officer's actions were justified. We find this to be the case, as shown in Table 8. Respondents were 13 percent (9 percentage points) less likely to say that the officer is morally responsible for the victim's death when the story specified that the victim was armed. Participants are also 19-22 percent less likely to agree that the officer should face penalties within the department or legally, respectively.

Table 8: Online experiment: perceptions of the police killing, depending on whether we specify that the victim had a weapon or not.

	Moral Reponsibility (1)	Department Penalty (2)	Legal Penalty (3)
Weapon	-0.09***	-0.75***	-0.83***
	(0.02)	(0.08)	(0.08)
Mean Dep. Var.	0.72	3.93	3.73
SD Dep. Var.	0.45	2.15	2.16
N	2402	2402	2402

Notes: The Weapon variable is equal to 1 if we specify that the victim was armed. The outcome in column 1 is a dummy equal to 1 if the respondents think that the police officer is morally responsible for the victim's death. The outcomes in columns 2 and 3 is support, on a scale from 1 to 7, for department and legal penalties, respectively. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for stories that do not specify if the victim had a weapon. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In addition to independent effects of narrative structure and the presence of a weapon on participants' responses, we also proposed a third hypothesis related to the interaction of the two treatments. In particular, we conjectured that obfuscatory narrative structure would be especially impactful when the story did not mention a weapon. This hypothesis was based on the idea that the presence of a weapon would lead some particants to determine that the shooting was justified regardless of how information was presented in the story. To examine this hypothesis, Table 9 breaks down the analysis presented in Table 7 by whether the story included a clause stating that the decedent "was reportedly armed" (Panel B) or not (Panel A).

In line with our hypothesis, the point estimates are greater in magnitude for *all 15 effects* related to the use of obfuscatory language reported in each panel when the story omits any mention of a weapon. In the absence of information about a weapon, the estimated effects are especially great when the story does not explicitly identify an agent (*No Agent* and *Intransitive*). The estimated effects are also negative for the *Passive* treatment in this case, but still mostly do not rise to statistical significance at conventional levels. Overall, the

Table 9: Online experiment: narrative structure and judgment of the event, by presence or absence of a weapon

		Moral Reponisbility			Department Penalty			Legal Penalty	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: No mention of victim weapon				
Obfuscation	-0.10***			-0.35**			-0.32**		
	(0.03)			(0.14)			(0.14)		
No Explicit Agent		-0.12***			-0.40***			-0.40***	
		(0.03)			(0.14)			(0.14)	
Passive		-0.06*	-0.06*		-0.24	-0.24		-0.16	-0.16
		(0.04)	(0.04)		(0.16)	(0.16)		(0.16)	(0.16)
No Agent + Nominalization			-0.08**			-0.37**			-0.40**
			(0.04)			(0.17)			(0.17)
Intrasitive + Nominalization			-0.15***			-0.44***			-0.41**
			(0.04)			(0.16)			(0.17)
Mean Dep. Var.	0.80	0.80	0.80	4.44	4.44	4.44	4.25	4.25	4.25
SD Dep. Var.	0.40	0.40	0.40	2.08	2.08	2.08	2.06	2.06	2.06
N	1201	1201	1201	1201	1201	1201	1201	1201	1201
					Panel B: Victim has a weapon				
Obfuscation	-0.04			-0.12			-0.15		
	(0.03)			(0.13)			(0.14)		
No Explicit Agent		-0.07**			-0.23			-0.26*	
		(0.03)			(0.14)			(0.14)	
Passive		0.01	0.01		0.11	0.11		0.08	0.08
		(0.04)	(0.04)		(0.17)	(0.17)		(0.17)	(0.17)
No Agent + Nominalization			-0.05			-0.25			-0.28*
			(0.04)			(0.16)			(0.17)
Intrasitive + Nominalization			-0.08**			-0.21			-0.24
			(0.04)			(0.17)			(0.17)
Mean Dep. Var.	0.66	0.66	0.66	3.51	3.51	3.51	3.29	3.29	3.29
SD Dep. Var.	0.47	0.47	0.47	2.11	2.11	2.11	2.16	2.16	2.16
N	1201	1201	1201	1201	1201	1201	1201	1201	1201

Notes: Panel A, we presents results when we do not specify that the victim was armed, while we do in Panel B. Obfuscation is equal to 1 if the sentence structure is passive, no agent, or intransitive. No explicit agent is equal to 1 if the sentence structure is no agent or intransitive. The outcome in columns 1-3 is a dummy equal to 1 if the respondents think that the police officer is morally responsible for the victim's death. The outcomes in the remaining columns are support, on a scale from 1 to 7, for department penalties (columns 4-6) and legal penalties (columns 7-9), respectively. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

estimated effect of *Any Obfuscation* and *No Explicit Agent* are negative and statistically significant for all three outcomes when the story does not mention the potential presence of a weapon. In contrast, the estimated effects for *Any Obfsucation* are much smaller in magnitude and not statistically significant when the story states that the decedent was reportedly armed.

Appendix Table D.4 shows similar patterns for two additional outcomes: whether the respondents thought that the officer was justified in shooting the person; and whether they thought that the officer was depicted negatively in the story. We find that – especially for stories that do not mention a weapon – more obfuscation both increases the perception that the officer was justified in shooting and decreases feelings that the officer was negatively depicted.

Broader Perceptions of Policing Harms and Demand for Reform. We next investigate how narrative structure affects broader perceptions of policing harms and demand for reform, beyond this particular incident. Columns 1-3 of Table 10 report results for how narrative structure affects how much participants would donate to a non-profit focused on police reform (vs. officer well-being), while Columns 4-6 report analogous results for participants' estimate of the prevalence of police killings in the United States. The point estimates for donations are negative, but generally smaller in magnitude than those related to the participants' judgments about the specific event shown in Table 7. In this case, the narrative structure with *No Explicit Agent* reduces donations by about 4 percent (2.5 percentage points), while the use of *Passive* voice continues to have negligible effects. Interestingly, the use of more obfuscatory sentence structures also tends to reduce participants' estimates of the number of annual police killings in the United States. This suggests that some of the decline in support for reform may arise due to a reduced salience of police killings as a social issue when obfuscatory language is used in the story.

Informational Content. The use of obfuscatory sentence structures could potentially affect perceptions and judgments about the police killing described in our experiment in multiple ways. A natural distinction is whether the fundamental information about the incident that participants take away from the story is altered or whether the effects estimated above instead reflect the differential impact of the same information.[24] The latter

[24] That viewers may take away different information content is motivated by the fact that the information transmitted across the four treatment arms of the experiment is not equivalent. While the *Active* (`'the police officer killed the man'`) and *Passive* (`'the man was killed by the police officer'`) trans-

Table 10: Online experiment: narrative structure and perceptions of policing, for all cases

	Donation Reform			Yearly Police Killings		
	(1)	(2)	(3)	(4)	(5)	(6)
Obfuscation	-2.1			-89.9*		
	(1.4)			(52.7)		
No Explicit Agent		-2.9*			-134.1**	
		(1.5)			(55.9)	
Passive		-0.7	-0.7		-2.0	-2.0
		(1.8)	(1.8)		(64.7)	(64.7)
No Agent + Nominalization			-2.1			-163.3**
			(1.8)			(64.5)
Intrasitive + Nominalization			-3.6**			-104.3
			(1.8)			(64.9)
Mean Dep. Var.	67.47	67.47	67.47	1352.56	1352.56	1352.56
SD Dep. Var.	31.13	31.13	31.13	1164.26	1164.26	1164.26
N	2402	2402	2402	2402	2402	2402

Notes: The outcome in columns 1-3 is number of dollars out of their 100 dollar donation that respondents want to give to an organization supporting police reform. The outcome in columns 4-6 the estimated number of police killings each year. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

could result, for example, if more active language leads readers to develop a more vivid picture of what happened or evokes a stronger emotional response.

To shed some light on these possible mechanisms, we explore whether differences in the narrative structure affect how people recall and/or retell the story. At the end of the survey, we asked participants to retell the news story that they had read in their own words. For half of our sample, two research assistants, who were blind to the treatment arms, coded the content of the sentences to capture whether the person explicitly said that the police officer had killed or shot a person.[25]

Table 11: Online experiment: participants retelling of the story.

	Explicit Police Shooting (1)	Active Voice (2)	No Explicit Agent (3)
Passive	-0.01	-0.14***	0.02
	(0.03)	(0.04)	(0.03)
No Agent + Nominalization	-0.02	-0.27***	0.12***
	(0.03)	(0.04)	(0.03)
Intrasitive + Nominalization	-0.16***	-0.35***	0.22***
	(0.03)	(0.04)	(0.03)
Mean Dep. Var.	0.92	0.73	0.07
SD Dep. Var.	0.26	0.44	0.26
N	1198	1198	1198

Notes: We asked participants to write what they recalled of the news story. We hand-classifed the text to capture whether the person explicitly said that at police officer shot or killed the person (column 1); whether they used an active voice (column 2); and whether there was no agent in the retell of the killing (column 3). We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results are presented in Column 1 in Table 11. In the Active treatment arm, the vast majority of the respondents explicitly mentioned the role of the police officer in the killing (92 percent) in their re-telling of the story. Interestingly, this varied little across the *Active*, *Passive*, and *No Agent* treatments, all of which use a form of the word 'kill', but participants who read the *Intransitive* version of the story were 16 percentage points less

mit clearly who is the agent, the patient, and the action (albeit in a different order), the agent is not clear in the *No Agent* (`‘man was killed in an officer-involved shooting’`) and both the agent and the causal action are not fully specified in the *Intransitive* (`‘the man died in an officer-involved shooting’`).

[25]The two coders agreed on 85% of sentences. Our team recoded those that they disagreed on.

likely to identify the police officer as having killed or shot someone. This suggests that the use of *Intransitive* sentence structures may fundamentally alter the information that viewers take away from the story. Moreover, it also suggests that the impact of the *Passive* and *No Agent/Nominalization* sentence structures in the experiment is not likely related to the participants' understanding of the fundamental information in the story but is likely driven by the response to this information that the narrative structure evoked.

Language Transmission. We use the particpants' open-ended retelling of the story to explore a final question related to the impact of narrative structure: Does obfuscatory language get retransmitted as information is shared with others? In evaluating the survey responses to answer this question, we measure the respondents' use of obfuscatory sentence structures, presenting these results in Columns 2 and 3 of Table 11. Strikingly, the estimates reveal that people tend to mimic the sentence structures they read when retelling the story themselves several minutes later. As shown in Column 2, they are significantly less likely to use the active voice following treatments using any of the obfuscatory sentence structures. The effect sizes are increasing with our ex-ante expected gradient of obfuscation: i.e., the largest decrease in the use of active voice is generated in the *Intransitive* arm (-35 percentage points), followed by the *No Agent* (-27 percentage points) and *Passive* (-14 percentage points). Similarly, Column 3 shows that, when there was *No Explicit Agent* in the story itself, respondents tend to also use sentence structures that do not explicitly acknowledge the police officer as a cause of the killing (+12 percentage points for *No Agent* and +22 percentage points for *Intransitive*). Notice that these effects arise despite the fact that the phrase "officer-involved shooting" occurs twice within each of these stories, suggesting that many participants do not interpret this phrase to automatically imply that the officer was the shooter. While the implications of this analysis are certainly limited by the short-term nature of the recall exercise in the experiment, the results suggest that there may be broader spillover effects of obfuscatory language structures. They suggest, in particular, that the media use of obfuscation may not only affect viewers directly, but may also shape the information they subsequently pass on to others.

Overall, our online experiment shows that obfuscation influences both the read of the situation; broader perceptions on policy issues; and how a story is then retold. However, there is variation in the importance of obfuscation. Consistently across contexts, nominalization and not including an explicit agent influences all of our measures of perception of a police killing, and its social consequences. Using a passive instead of an active voice also matters, but less strongly, and not in all cases (for example, not when we specify that the victim was armed). In other words, the more tortured the language choice, the greater the impact it has on what people retain from a news story.

Table 12: Obfuscation in news of police killings, depending on whether the victim was allegedly armed or not

	All Sentences		1st Sentence	
Dimension of Obfuscation:	Any (1)	No Explicit Agent (2)	Any (3)	No Explicit Agent (4)
	Panel A: Unknown or No Reported Weapon for Victim of Police Shooting			
Police Killing	0.130*** (0.011)	0.040*** (0.007)	0.175*** (0.013)	0.085*** (0.009)
Observations	201,780	201,780	79,292	79,292
Mean Civ.Shoot.	0.2935	0.1354	0.2814	0.1266
	Panel B: Reported Weapon for Victim of Police Shooting			
Police Killing	0.058*** (0.008)	0.024*** (0.005)	0.104*** (0.010)	0.055*** (0.006)
Observations	411,450	411,450	158,588	158,588
Mean Civ.Shoot.	0.2935	0.1354	0.2814	0.1266
Controls	Story Controls+DMA FE			

Notes: Standard errors clustered by subject (*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$). This table breaks out differential obfuscation for stories about police killings, relative to stories about civilian killings, depending on whether the victim was allegedly armed or not, as defined by MPV. Panel A is for cases where the victim was allegedly unarmed, while Panel B is for cases where the victim was allegedly armed. The analysis is at the sentence level. Our sample includes incidents and news stories where a suspect was identified for civilian killings. We define 'obfuscation' as having passive forms, no agent, intransitive verbs, or nominalization. We define 'No Explicit Agent' as having no agent or intransitive verbs. See Section 2 for more details. Source: News Exposure.

5.3 Differential Obfuscation

Our experimental results showed that for any given case of police killings, obfuscation is most effective in changing attitudes of the audience when the victim is not reported to have a weapon. We explore the extent of differential obfuscation in the news coverage across this dimension, and test if indeed there is higher obfuscation in the cases in which it matters most.

The Mapping Police Violence data includes a variable that describes whether the victim was allegedly armed or not. Using this variable, we break out stories about police killings depending on whether the victim was allegedly armed or not. Results are presented in Table 12. The top panel is for cases where the victim was not armed, and the bottom panel is for cases where the victim was allegedly armed. Panel A shows that there is indeed obfuscation happening for cases where the victim was not armed. Comparing the coefficients in Panel A to those in Panel B shows that, in fact, there is twice as much obfuscation when the victim is unarmed than allegedly armed. For all the pairwise comparison across the coefficients, the estimates for 'no weapon' are larger and statistically different from those with 'weapon'[26]. These are likely to be the kinds of cases where obfuscation might be especially favorable for police officers, given results from our lab experiments.

Another case in which obfuscation might be more favorable for the police officers is when there is body-worn camera footage. We test if there is differential obfuscation across this dimension, and present the results in the Table D.8. The MPV dataset once more provides a flag for such a variable that we use in the analysis. The top panel compares civilian killings with police killings where there was body-worn camera footage, and the bottom panel is for cases in which there was not. Once more, comparing the coefficients in Panel A to those in Panel B shows that there is more obfuscation when there was body-worn camera footage. The difference of magnitude is again almost twice as large when there is body-worn camera footage as when there is not. For all the pairwise comparisons across coefficients, the estimates for 'body-worn camera' are larger and statistically different from those with 'no-body-worn camera' at the 5% level.[27]

6 Conclusion

The main aim of this paper is to provide new empirical and experimental evidence on two interrelated questions that have received a great deal of attention in recent years: whether there is systematic obfuscation of responsibility in media coverage of police killings and, if so, whether this matters for perceptions of both the incident in question and the potential harms from policing more generally. To answer these questions, we collected comprehensive data on police killings, civilian homicides, and television news coverage in the United

[26]The p-value is less than 0.01 for all cases, except *No-Explicit Agent* in the any obfuscation comparison, which has a p-value of 0.03

[27]In Appendix Figure C.3, we present heterogeneity results by different dimensions of the DMA or station characteristics (political leaning, size, reach of the station), and by incident characteristics (sex and race of the victim, virality of the story). We do not find any clear variation in obfuscation along these dimensions, suggesting that this phenomenon is commonplace across different contexts.

States covering the period 2013-19 and also conducted an online experiment in which we varied the narrative structure used to report a story about a police killing.

The results of our analyses provide clear and robust evidence on both questions. First, we document that there is more obfuscation when an officer was responsible for the killing relative to civilian homicides for which the media faced a similar choice of language. The use of obfuscatory language is especially common in the story's lead sentence, which is most salient to viewers. Second, we show that obfuscation matters: in the experiment, respondents' read of the situation varies with the degree of obfuscation. They are less likely to think that the officer is morally responsible and to ask for penalties when there is obfuscation – all the more so when we do not specify that the civilian was armed. Prompted by the experimental results, we close the paper by asking a third question: whether there is differential obfuscation in cases where it might especially benefit the police – e.g., cases when the victim was not armed or when body camera footage is available. We find a doubling of the use of obfuscation in these cases.

Our results also indicate that the narrative structures employed by media outlets, which often mirror those used in press releases and tweets by police departments and unions, impact the way that the public understands the harms from policing more generally, as well as support for police accountability and reform. These broader effects of language are important given the growing discussion on policy changes and reforms society might implement to improve public safety (Akbar (2020); Bursztyn *et al.* (2022b)) in light of the significant negative externalities of police violence on cities and individuals (DiPasquale & Glaeser (1998); Ang (2020); González & Prem (2022)). One question our analysis does not answer is why the media is especially likely to use obfuscatory language for police killings. Documenting when terms such as 'officer-involved' began to be widely used by the media and whether, for example, the use of obfuscatory language in police press conferences and press releases spills over directly to the language used by the local media are promising avenues for future research.

Finally, while our analysis focuses on the semantic structure of language in the context of the media's coverage of police killings, our approach, studying the scale and consequences of obfuscation, offers a practical and widely applicable template for other topics covered in news outlets or social media. For example, how does media cover different forms of interpersonal violence, and how does this influence perceptions of responsibility and support for policy change? Beyond the crime and criminal justice space, the way in which stories are told might also matter for many other economic and social issues, such as income inequality, immigration, climate change or health. This analysis could also easily be extended to study the language structures used by a wider set of actors - such as political speech, corporate messaging, and social media influencers - broadening our focus on traditional media.

References

Akbar, Amna A. 2020. An Abolitionist Horizon for (Police) Reform. *California Law Review*, **108**, 1781.

Alesina, Alberto, Ferroni, Matteo F, & Stantcheva, Stefanie. 2021. *Perceptions of Racial Gaps, their Causes, and Ways to Reduce Them*. Tech. rept.

Andre, Peter, Haaland, Ingar, Roth, Chrisotpher, & Wohlfart, Johannes. 2021. Inflation narratives.

Ang, Desmond. 2020. The Effects of Police Violence on Inner-City Students. *The Quarterly Journal of Economics*, 09.

AP. 2020. *The Associated Press Stylebook 2020-2022*. The Associated Press.

Baker, Scott R., Bloom, Nicholas, & Davis, Steven J. 2016. Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, **131**(4), 1593–1636.

Balko, Radley. 2014. The curious grammar of police shootings. *The Washington Post*.

Blachor, Devorah. 2020. How to use the Past Exonerative Tense to Uphold White Supremacy. *McSweeney's*.

Bohner, Gerd. 2001. Writing about rape: Use of the passive voice and other distancing text features as an expression of perceived responsibility of the victim. *British Journal of Social Psychology*, **40**(4), 515–529.

Bonial, Claire, Babko-Malaya, Olga, Choi, Jinho D, Hwang, Jena, & Palmer, Martha. 2010. Propbank annotation guidelines. *Center for Computational Language and Education Research, CU-Boulder*, **9**.

Bursztyn, Leonardo, Egorov, Georgy, Haaland, Ingar, Rao, Aakaash, & Roth, Christopher. 2022a. *Justifying Dissent*. Tech. rept. 141.

Bursztyn, Leonardo, Rao, Aakaash, Roth, Christopher P, & Yanagizawa-Drott, David H. 2022b. *Misinformation During a Pandemic*. Working Paper. National Bureau of Economic Research.

Cagé, Julia. 2020. Media competition, information provision and political participation:

Evidence from French local newspapers and elections, 1944–2014. *Journal of Public Economics*, **185**, 104077.

Chan, Eugene Y, & Maglio, Sam J. 2020. The voice of cognition: Active and passive voice influence distance and construal. *Personality and Social Psychology Bulletin*, **46**(4), 547–558.

Chestnut, Eleanor K, & Markman, Ellen M. 2018. "Girls are as good as boys at math" implies that boys are probably better: A study of expressions of gender equality. *Cognitive science*, **42**(7), 2229–2249.

Chiang, Chun-Fang, & Knight, Brian. 2011. Media Bias and Influence: Evidence from Newspaper Endorsements. *The Review of Economic Studies*, **78**(3), 795–820.

Conner, Andrew, Azrael, Deborah, Lyons, Vivian H, Barber, Catherine, & Miller, Matthew. 2019. Validating the National Violent Death Reporting System as a source of data on fatal shootings of civilians by law enforcement officers. *American journal of public health*, **109**(4), 578–584.

Couttenier, Mathieu, Hatte, Sophie, Thoenig, Mathias, & Vlachos, Stephanos. 2021. Anti-Muslim Voting and Media Coverage of Immigrant Crimes. *The Review of Economics and Statistics*, 1–33.

Dahl, Gordon, & DellaVigna, Stefano. 2009. Does movie violence increase violent crime? *The Quarterly Journal of Economics*, **124**(2), 677–734.

De Freitas, Julian, DeScioli, Peter, Nemirow, Jason, Massenkoff, Maxim, & Pinker, Steven. 2017. Kill or die: Moral judgment alters linguistic coding of causality. *J. Exp. Psychol. Learn. Mem. Cogn.*, **43**(8), 1173–1182.

DellaVigna, Stefano, & Kaplan, Ethan. 2007. The Fox News Effect: Media Bias and Voting*. *The Quarterly Journal of Economics*, **122**(3), 1187–1234.

DellaVigna, Stefano, & La Ferrara, Eliana. 2015. Economic and social impacts of the media. *Pages 723–768 of: Handbook of media economics*, vol. 1. Elsevier.

DellaVigna, Stefano, Enikolopov, Ruben, Mironova, Vera, Petrova, Maria, & Zhuravskaya, Ekaterina. 2014. Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics*, **6**(3), 103–32.

DiPasquale, Denise, & Glaeser, Edward L. 1998. The Los Angeles Riot and the Economics of Urban Unrest. *Journal of Urban Economics*, **43**(1), 52–78.

Dixon, Travis L, & Williams, Charlotte L. 2015. The changing misrepresentation of race and crime on network and cable news. *Journal of Communication*, **65**(1), 24–39.

Djourelova, Milena. 2021. *Media Persuasion through Slanted Language: Evidence from the*

Coverage of Immigration. Tech. rept.

Durante, Ruben, & Zhuravskaya, Ekaterina. 2018. Attack when the world is not watching? US news and the Israeli-Palestinian conflict. *Journal of Political Economy*, **126**(3), 1085–1133.

Duxbury, Scott W, Frizzell, Laura C, & Lindsay, Sadé L. 2018. Mental illness, the media, and the moral politics of mass violence: The role of race in mass shootings coverage. *Journal of Research in Crime and Delinquency*, **55**(6), 766–797.

Eisensee, Thomas, & Strömberg, David. 2007. News Droughts, News Floods, and U. S. Disaster Relief. *The Quarterly Journal of Economics*, **122**(2), 693–728.

Engle, Robert F, Giglio, Stefano, Kelly, Bryan, Lee, Heebum, & Stroebel, Johannes. 2020. Hedging Climate Change News. *The Review of Financial Studies*, **33**(3), 1184–1216.

Enikolopov, Ruben, Petrova, Maria, & Zhuravskaya, Ekaterina. 2011. Media and political persuasion: Evidence from Russia. *American Economic Review*, **101**(7), 3253–85.

Frazer, Alexandra K, & Miller, Michelle D. 2009. Double standards in sentence structure: Passive voice in narratives describing domestic violence. *Journal of language and social psychology*, **28**(1), 62–71.

Gardner, Matt, Grus, Joel, Neumann, Mark, Tafjord, Oyvind, Dasigi, Pradeep, Liu, Nelson F., Peters, Matthew, Schmitz, Michael, & Zettlemoyer, Luke S. 2017. AllenNLP: A Deep Semantic Natural Language Processing Platform.

Gay, Victor, Hicks, Daniel L, Santacreu-Vasut, Estefania, & Shoham, Amir. 2018. Decomposing culture: an analysis of gender, language, and labor supply in the household. *Review of Economics of the Household*, **16**(4), 879–909.

Gentzkow, Matthew, & Shapiro, Jesse M. 2008. Preschool television viewing and adolescent test scores: Historical evidence from the Coleman study. *The Quarterly Journal of Economics*, **123**(1), 279–323.

Gentzkow, Matthew, & Shapiro, Jesse M. 2010. What drives media slant? Evidence from US daily newspapers. *Econometrica*, **78**(1), 35–71.

Gentzkow, Matthew, Kelly, Bryan, & Taddy, Matt. 2019. Text as Data. *Journal of Economic Literature*, **57**(3), 535–74.

Gentzkow, Matthew A., & Shapiro, Jesse M. 2004. Media, Education and Anti-Americanism in the Muslim World. *Journal of Economic Perspectives*, **18**(3), 117–133.

Giglio, Stefano, Kelly, Bryan, & Stroebel, Johannes. 2021. Climate Finance. *Annual Review of Financial Economics*, **13**(1), 15–36.

Gilliam, Franklin D Jr, & Iyengar, Shanto. 2000. Prime suspects: The influence of local television news on the viewing public. *American Journal of Political Science*, 560–573.

Gilliam Jr, Franklin D, Iyengar, Shanto, Simon, Adam, & Wright, Oliver. 1996. Crime in black and white: The violent, scary world of local news. *Harvard International Journal of press/politics*, **1**(3), 6–23.

González, Felipe, & Prem, Mounu. 2022. Police Violence, Student Protests, and Educational Performance. *The Review of Economics and Statistics*, 1–46.

Grimmer, Justin. 2010. A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis*, **18**(1), 1–35.

Grosjean, Pauline, Masera, Federico, & Yousaf, Hasin. 2022. Inflammatory Political Campaigns and Racial Bias in Policing.

Grunwald, Ben, Nyarko, Julian, & Rappaport, John. 2022. *Police agencies on Facebook overreport on Black suspects*. Working Paper.

Hassan, Tarek A, Hollander, Stephan, van Lent, Laurence, & Tahoun, Ahmed. 2019. Firm-Level Political Risk: Measurement and Effects*. *The Quarterly Journal of Economics*, **134**(4), 2135–2202.

Henley, Nancy M, Miller, Michelle, & Beazley, Jo Anne. 1995. Syntax, semantics, and sexual violence: Agency and the passive voice. *Journal of Language and Social Psychology*, **14**(1-2), 60–84.

Humphries, John Eric, Mader, Nicholas S, Tannenbaum, Daniel I, & van Dijk, Winnie L. 2019 (Aug.). *Does Eviction Cause Poverty? Quasi-Experimental Evidence from Cook County, IL*.

Imai, Kosuke, & Khanna, Kabir. 2016. Improving ecological inference by predicting individual ethnicity from voter registration records. *Political Analysis*, 263–272.

Jakiela, Pamela, & Ozier, Owen. 2018. Gendered language. *World Bank Policy Research Working Paper*.

Kazenin, Konstantin I. 2001. *67. The passive voice*.

Kearney, Melissa S., & Levine, Phillip B. 2015. Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing. *American Economic Review*, **105**(12), 3597–3632.

Khanna, K, Imai, K, & Hubert, J. 2017. *Who are You? Bayesian Prediction of Racial Category Using Surname and Geolocation*. Tech. rept. Technical report, The Comprehensive R Archive Network.

Lee, Kenton, He, Luheng, & Zettlemoyer, L. 2018. Higher-order Coreference Resolution with Coarse-to-fine Inference. *In: NAACL-HLT*.

Lussos, Rachael Graham, & Fernandez, Lourdes. 2018. Assault and accusation without agents: Verb voice in media narratives of campus sexual assault. *Journal of Mason Graduate Research*, **5**(2), 108–127.

Martin, Gregory J, & McCrain, Joshua. 2019. Local news and national politics. *American Political Science Review*, **113**(2), 372–384.

Martin, Gregory J, & Yurukoglu, Ali. 2017. Bias in cable news: Persuasion and polarization. *American Economic Review*, **107**(9), 2565–99.

Mastrorocco, Nicola, & Minale, Luigi. 2018. News media and crime perceptions: Evidence from a natural experiment. *Journal of Public Economics*, **165**, 230–255.

Mastrorocco, Nicola, & Ornaghi, Arianna. 2020. *Who Watches the Watchmen? Local News and Police Behavior in the United States*. Tech. rept.

Misra, Rishabh. 2018 (06). *News Category Dataset*.

Moreno-Medina, Jonathan. 2021. *Local Crime News: Extent, Causes and Consequences*. Tech. rept.

Niemi, Laura, & Young, Liane. 2016. When and why we see victims as responsible: The impact of ideology on attitudes toward victims. *Personality and social psychology bulletin*, **42**(9), 1227–1242.

Northcutt Bohmert, Miriam, Allison, Kayla, & Ducate, Caitlin. 2019. "A rape was reported": construction of crime in a university newspaper. *Feminist Media Studies*, **19**(6), 873–889.

Philippe, Arnaud, & Ouss, Aurélie. 2018. "No Hatred or Malice, Fear or Affection": Media and Sentencing. *Journal of Political Economy*, **126**(5), 2134–2178.

Pinker, Steven. 2007. *The stuff of thought: Language as a window into human nature*. Penguin.

Sanh, Victor, Debut, Lysandre, Chaumond, Julien, & Wolf, Thomas. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Shi, Peng, & Lin, Jimmy. 2019. Simple BERT Models for Relation Extraction and Semantic Role Labeling. *ArXiv*, **abs/1904.05255**.

Toolan, Michael J. 2013. *Narrative: A critical linguistic introduction*. Routledge.

Trew, Tony. 1979. Theory and ideology at work. *Pages 94–116 of: Language and control*. Routledge.

Wolff, Phillip. 2003. Direct causation in the linguistic coding and individuation of causal events. *Cognition*, **88**(1), 1–48.

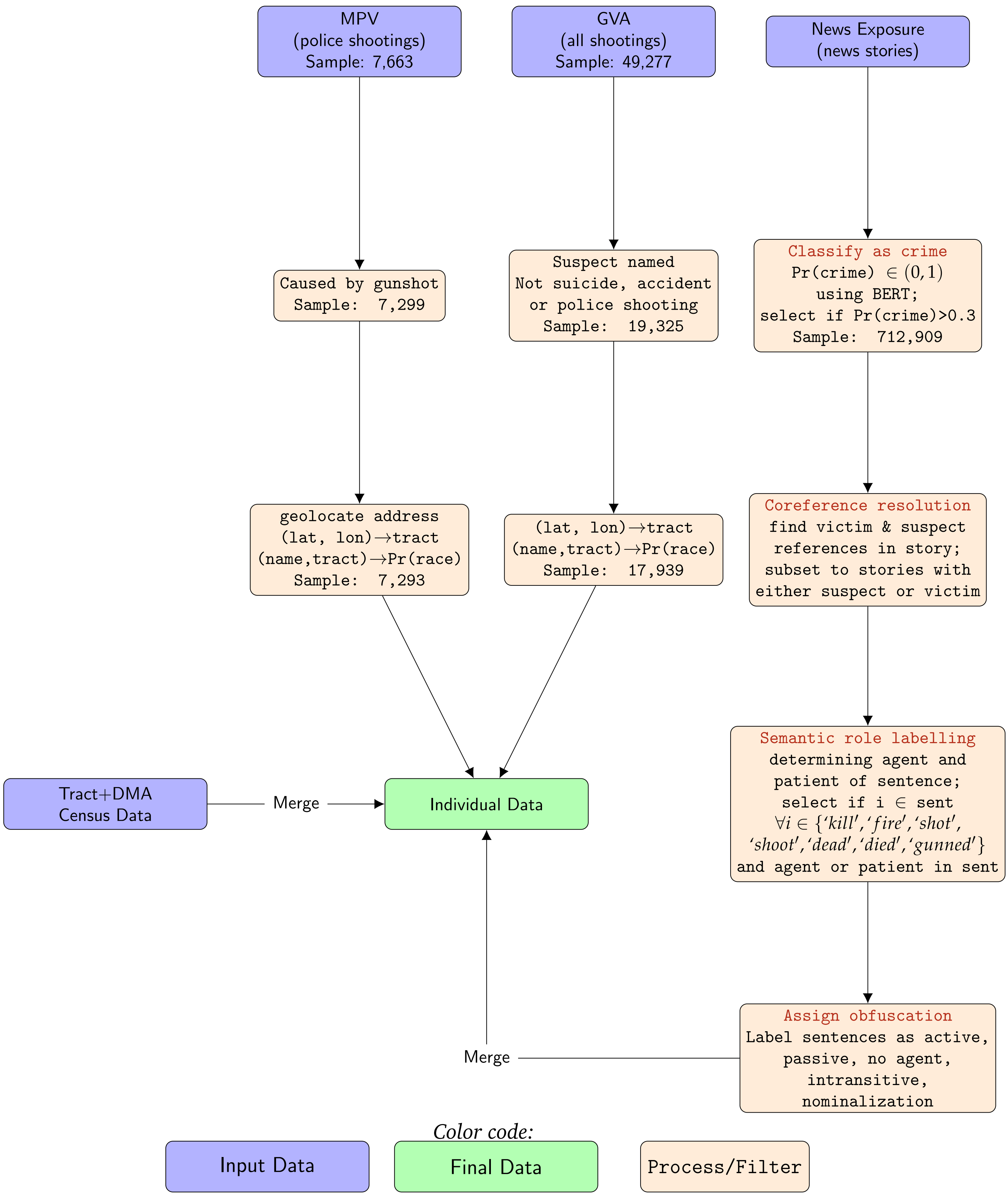
Yasmin, Musarat. 2021. Asymmetrical gendered crime reporting and its influence on readers: A case study of Pakistani English newspapers. *Heliyon*, **7**(8), e07862.

Appendix A Data processing

This Appendix describes in detail the data processing from the raw data from the MPV, GVA and News Exposure datasets, to the final data used for text analysis. First, we present in figure A a flowchart of our data cleaning steps. Second, in sections A.1 - A.4, we describe in turn each of our data processing steps:

1. Identifying news stories about shootings (section A.1)
2. Co-reference resolution (section A.2)
3. Semantic role labeling (section A.3)
4. Assigning degree of obfuscation to each sentence and story (section A.4)

Figure A.1: Data processing



A.1 Classification of a story as crime-related

This section briefly describes how we classify each matched text from the TV caption data as about crime or not. We employ a state-of-the-art Natural Language Processing model, based on *BERT* (Bidirectional Encoder Representations from Transformers).[28] *BERT* is very accurate in understanding contextual embeddings for words. This is important in our context because many non-crime stories could include words typically used in crime stories. One example would be the use of the word 'shot' to describe a soccer or basketball action instead of the action of a gun.

We employ the pre-trained smaller and faster version of *BERT*, called *DistilBERT* (Sanh *et al.*, 2019). We re-train it to classify stories across '**crime**' and '**not crime**' categories. For this latter task, we need a database that accurately labels news stories on crime. We used an already labeled dataset of close to 200,000 short news stories of the Huffington Post, run between 2012 and 2018, collected and shared by Misra (2018). These data contains labels for **crime** news (3,405), as well as 20 other categories. We group together all other categories into the '**non-crime**' category. With this trained model, we estimate the probability that the matched caption is about crime or not. Manual inspection of a subsample of the predictions revealed that 99 percent of the stories with an estimated probability of being about crime of 0.3 or higher to indeed be about crime. This is the threshold we used for our classification. With this classification we go from 1'839.853 stories (1'170.573 from non-police killings; 669.280 from police killings) to 712,909 stories (426,728 from non-police killings; 286,181 from police killings).

Figure A.2 presents the histogram of the probabilities across police and non-police killings.

[28] *BERT* (Bidirectional Encoder Representations from Transformers) is a neural network model of language that has proven to be incredibly successful among a host of tasks in NLP. There are several technical features in BERT, but perhaps the most important is that it trains the model not only looking using previous words in a text, but also future ones. The standard model allows up to 512 words (or tokens) in a text. That means that for cases in which our stories have more than 512 tokens in a text, we truncate the length of the text at 512 and drop all the remaining tokens. The network has 7 layers, and it works with a type of word embedder model that captures the context in which the word is being used. Since 2019, *Google Search* announced that they had started applying BERT models for English language search queries within the US.

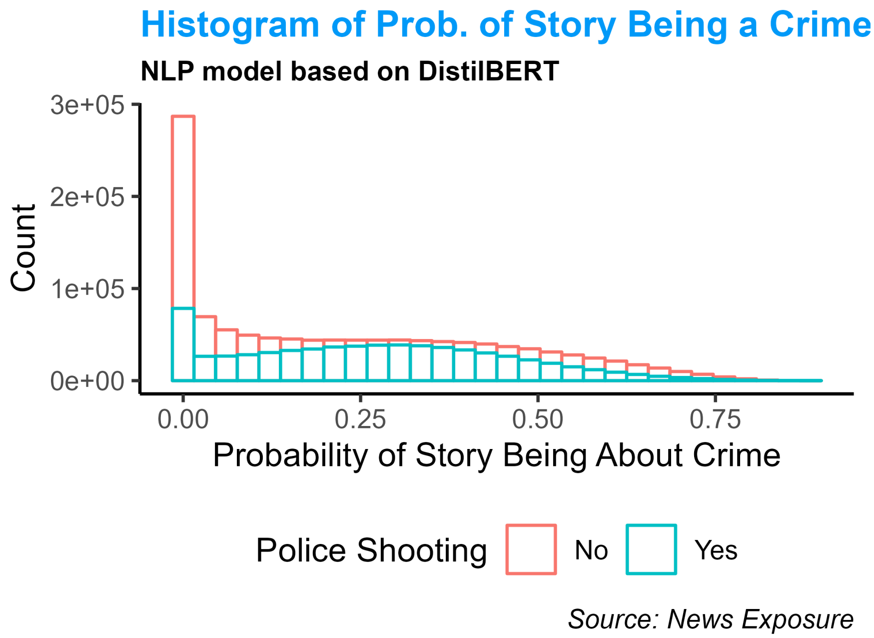


Figure A.2: Histogram of Probability of Story Being about Crime

A.2 Coreference resolution

A.2.1 Basic coreference resolution

Coreference resolution is the task of finding all expressions that refer to the same entity in a text. This is important, as in sentences where 'him' or 'man' is used, we want to know if it refers to the same individual (victim or perpetrator) or not.

We use the model proposed by Lee *et al.* (2018), with the library provided by Gardner *et al.* (2017). Here is an example of how the library works. Consider the following text in our database, which contains a story on a police killing:[29]

> Police say [victim's name] fled before officers shot him, and say officers found the rifle in a nearby apartment. We've learned [victim's name] had several run-ins with the law. In 2007, he pleaded guilty to possessing a firearm. [victim's name] was released from prison more than a month after violating parole. He is married with six children. Their youngest is 5 months old. [victim's brother name] says his brother [victim's first name] was shot in the head,

[29]We have replaced the name of the victim in the text.

and the bullet went through his cheek and hit his spinal cord. Police are still investigating. No officers were injured in the shooting. Heather Hope 17 news.

The output of the algorithm is a set of 'clusters' describing all the tokens recognized to be describing the same entity. Figure **??** presents the output for this example.

This example is define by the following 'spans' (or sets of tokens)[30]: ['[Victim's Name]', 'HIM', '[Victim's Name]','HE','[Victim's Name]','HE','[Victim's Brother Name]', 'HIS', 'HIS BROTHER [Victim's First Name]', 'HIS', 'HIS'].

Notice that the algorithm did a very good job identifying all the instances in which the *victim* ([Victim's Name]) is referenced in the text. There are only 2 small mistakes with the output: '[Victim's Brother Name]' and the immediately next 'HIS' is not referring to the victim, but his brother. In our case this is not particularly worrisome, as if we focus on sentences with a described victim getting being killed (which we describe in the Section A.3), we will correctly identify all instances in which the victim was described as being shot.

A.2.2 Using coreference resolution to identify relevant story

As the previous example elucidates, getting to the relevant part of the story we are interested is key. We use coreference resolution to identify the relevant story that we care about in the following way:

1. Find all coreferences in raw text.

2. Replace all references by 'Victim' if any span in the cluster includes the name of the victim.

3. Replace all references by 'Perpetrator' (how we will call the suspect - not including police) if any span in the cluster includes the name of the suspect.

4. Divide the text into sentences.

5. Define the story as all sentences between the first and last mention of Victim or Suspect.

[30]This is not quite exact. Actually, the algorithm provides the spans that capture the same entity with coordinates in terms of characters. So, for the first mention of '[Victim's Name]', the algorithm would show (11,22), which are the position of the first 'A' in '[Victim's Name]' and 22, which is the last 'A' in the same span.

Figure A.3: Example of coreference resolution

We illustrate the above process with another example.[31] Consider this raw text, which includes the story we are looking for (victim's references are **BOLDED**), and some other text which is not about the victim / shooting (in *italics*):

> *What the proud parents are saying about having their bundle of joy at the start of 20 - 13. [a6]open 5pm 365 - deko 2.* **A BAKERSFELD MAN** is in critical condition as a result of an officer involved shooting that happened just after the new year arrived. The **MAN'S** family claims Bakersfield police did not have to shoot 26 year - old **[VICTIM'S NAME]** who police say had a rifle and refused to put it down after being told to do so. The shooting happened at an apartment complex on the 7 hundred block of terrace way. Police say they were responding to calls of shots fired shortly after midnight. When they got there, police claim they saw **[VICTIM'S NAME]** holding a rifle and told **HIM** to drop it.

It will become the following text after steps 1-3:

> *What the proud parents are saying about having their bundle of joy at the start of 20 - 13. [a6]open 5pm 365 - deko 2* **VICTIM** is in critical condition as a result of an officer involved shooting that happened just after the new year arrived. **VICTIM'S** family claims Bakersfield police did not have to shoot 26 year - old **VICTIM** who police say had a rifle and refused to put it down after being told to do so. The shooting happened at an apartment complex on the 7 hundred block of terrace way. Police say they were responding to calls of shots fired shortly after midnight. When they got there, police claim they saw **VICTIM** holding a rifle and told **VICTIM** to drop it.

and finally, after steps 4-5, we define the story to be:

> **VICTIM** is in critical condition as a result of an officer involved shooting that happened just after the new year arrived. **VICTIM'S** family claims Bakersfield police did not have to shoot 26 year - old **VICTIM** who police say had a rifle and refused to put it down after being told to do so. The shooting happened at an apartment complex on the 7 hundred block of terrace way. Police say they

[31] Steps 2 and 3 above are not quite exact. We replace the reference to the victim for the name 'Pete' (same starting letter as Patient), and to the perpetrator for 'Adam' (same starting letter as Agent). We do this as the semantic role labeling model has been trained with human names, which makes the prediction explained in Section A.3 more accurate.

were responding to calls of shots fired shortly after midnight. When they got there, police claim they saw **VICTIM** holding a rifle and told **VICTIM** to drop it.

Once we have defined the story, and replaced all the instances in which the victim or suspect is being referred to in the text by the explicit 'Victim' and 'Perpetrator' (respectively), we can proceed to uncover who is the agent or the patient in each sentence reporting on the killing of the victim.

A.3 Semantic role labeling

In this section we present how we identify who (agent) does what (verb) to whom (patient) - the objective of semantic role labeling. We select for the analysis sentences within each story (as defined in Section A.2.2) that mentions the Victim or Perpetrator, and involves one of the relevant verbs or adverb modifiers we are interested in: 'kill', 'fire', 'shoot', 'dead', 'die' or 'gunned'.

For this task we use another BERT-type model proposed by Shi & Lin (2019). The output uses the PropBank annotation dataset (Bonial *et al.*, 2010). The model takes each arguments of each predicate in the sentence, and annotates it with their semantic roles in relation to the predicate. Each verb is defined to possibly have several type of predicates, such as:

- ARG-0 is usually PROTO-AGENT (who executes the verb)
- ARG-1 is usually PROTO-PATIENT (who gets affected by the action)
- ARG-2 is usually benefactive, instrument, attribute
- ARG-3 is usually start point, benefactive, instrument, attribute
- ARG-4 is usually end point (e.g., for move or push style verbs)

The following illustrates how the algorithm works in out case. Consider the following sentence:

> According to authorities, [Perpetrator] shot [Victim] in the left side of [Victim] head, killing [Victim] instantly.[32]

Figure A.4: Example of Semantic Role Labeling

The output from the model is presented in Figure A.4. It shows that the model identifies 3 actions or verbs ('according', 'shot', and 'killing'), and identifies if there is an agent or patient for that verb, plus other information related to that action.

Once we have this output, we can assign the category for each sentence in our sample. We describe this process next in Section A.4.

A.4 Assignment of obfuscation dimension for each sentence and story

Based on the linguistic framework layed out in Section **??**, we define the following classification for each sentence.

Passive: We categorize the sentence as *Passive* if any of the following transitive verbs appear in the sentence with the patient being 'Victim' (based on Section A.3), and the passive structure is being used. [33]

List of transitive verbs for killings: Kill, Shot, Gun, Murder, Shoot, Hit, Fire, Open (fire), Strike.

[32] We remind the reader that at this point we have replaced all instances of reference to either the victim by Pete, and the suspect by Adam. This is the coreference resolution task as described in Section A.2.

[33] The passive is relatively easy to identify based on the use of auxiliary verb be(such as 'be', 'was', 'were', 'are', etc) followed by the past participle of the main verb.

No Agent: We subset those sentences classified as passive, and classify them as *No Agent* if the agent identified is either empty of different from the known one (in civilian killings it is 'Perpetrator', while in police killings it is either 'officer', 'deputy', 'sherif', 'sergeant', 'detective', 'they' or 'SWAT' or slight modifications of these words).

Intrasitive: We classify the sentence as *Intransitive* if the following intransitive verbs in the sentence with the patient being 'Victim' (based on Section A.3).

List of intransitive verbs for killings: Die, (is) dead, (declared/found/pronounced) dead.[34]

Nominalization: We classify the sentence as being a nominalization if it includes a description of a shooting in the form of 'X involved shooting', 'X related shooting', 'shooting (death) of' or 'shooting (killing) of'.[35]

Lastly, the classification at the story level is equal to 1 if it includes any of the sentences above described for the respective obfuscation category.

Appendix B Online Experiment: Additional Materials

B.1 Treatment arms

Our experiment has 8 treatment arms: intransitive / active / passive / no-agent, interacted with whether the story mentions that the civilian killed was armed or not. We present each sentence structure in turn.

Active

A police officer killed a 52-year-old man on Friday night.

According to the Police Department, an officer responded to a home near 21st Street and Avenue C for a report of domestic violence just before 9:30 p.m. As the officer arrived, he came into contact with a 52-year old man [who was reportedly armed]. The police officer shot the man. The man was taken to the hospital, where he later died. No officer was hurt in the incident.

[34] Although declare, find and pronounce are transitive verbs, the agent of killing is not being explicitely acknowledged, and thus it effectively functions as the other verbs in this category.

[35] The '(death)' and '(killing)' in the parenthesis may or not appear in the text.

Passive

A 52-year-old man was killed by a police officer on Friday night.

According to the Police Department, an officer responded to a home near 21st Street and Avenue C for a report of domestic violence just before 9:30 p.m. As the officer arrived, he came into contact with a 52-year old man [who was reportedly armed]. The man was shot by police officers. The man was taken to the hospital, where he later died. No officer was hurt in the incident.

No agent

A 52-year-old man was killed in an officer-involved shooting on Friday night.

According to the Police Department, an officer responded to a home near 21st Street and Avenue C for a report of domestic violence just before 9:30 p.m. As the officer arrived, he came into contact with a 52-year old man [who was reportedly armed]. The man was shot in an officer-involved shooting. The man was taken to the hospital, where he later died. No officer was hurt in the incident.

Intransitive

A 52-year-old man died in an officer-involved shooting on Friday night.

According to the Police Department, an officer responded to a home near 21st Street and Avenue C for a report of domestic violence just before 9:30 p.m. As the officer arrived, he came into contact with a 52-year old man [who was reportedly armed]. The man got wounded in an officer-involved shooting. The man was taken to the hospital where he later died. No officer was hurt in the incident.

B.2 Experiment questions

1. Is the officer morally responsible for the man's death? (Yes / No)

2. On a scale from 1 to 7, how much do you agree with these statements? (7 being the strongest and 1 being the weakest):

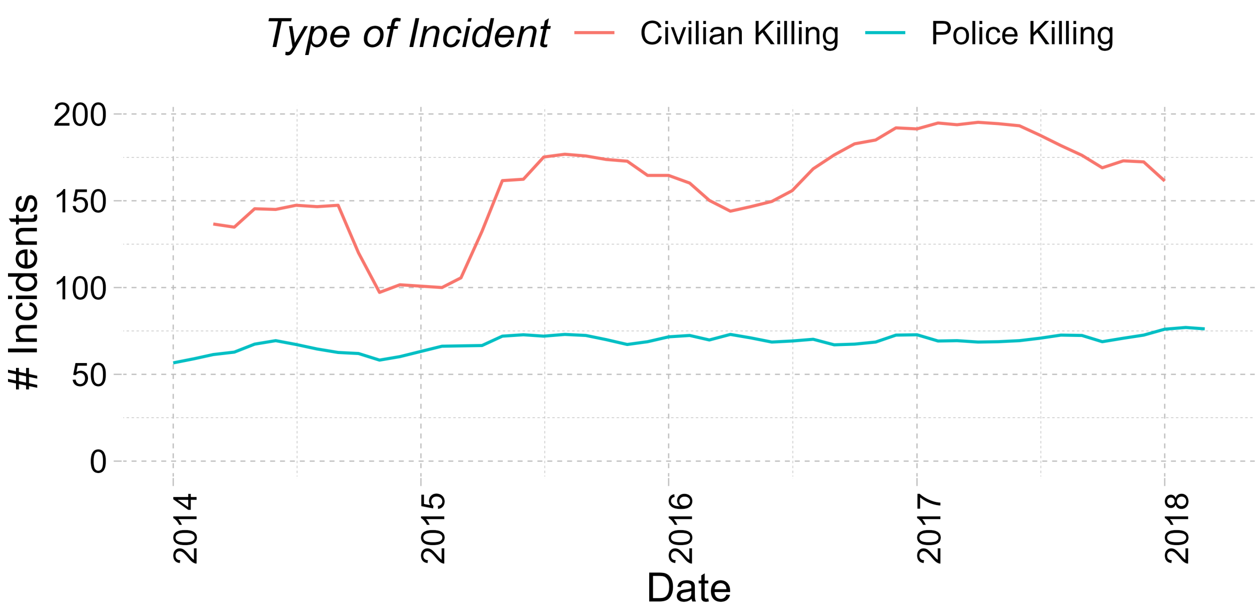
    - The police officer should face some penalties in their department for their actions
    - The police officer should face some legal penalties

3. One in 100 participants in this study will have the opportunity to donate $100 to an organization. This is in addition to their payment for participating in the study. Please choose how you want to split your donation among the two organizations below. If you are randomly selected, we will make an anonymous donation to each organization as you have decided below. (Answers need to add up to 100%)

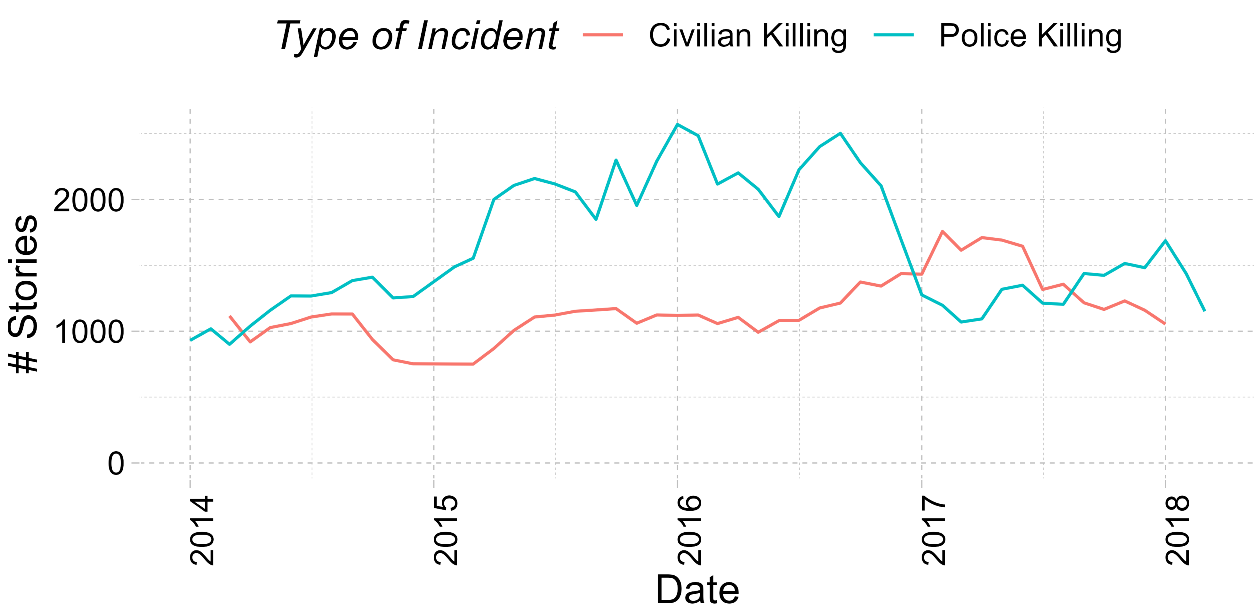
    - An organization that aims to improve officer safety as well as health and wellness in police
    - An organization that advocates to reform the police by increasing accountability, for example through officer training.

Appendix C Additional Figures

Figure C.1: Number of incidents and stories in sample



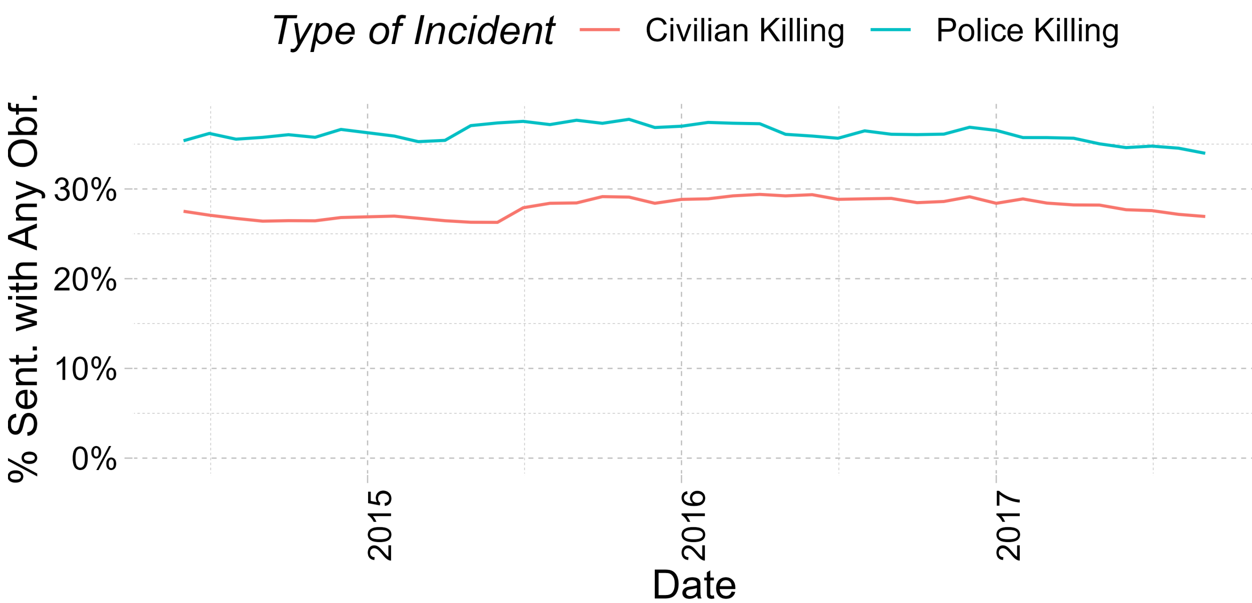
(a) Number of incidents



(b) Number of stories

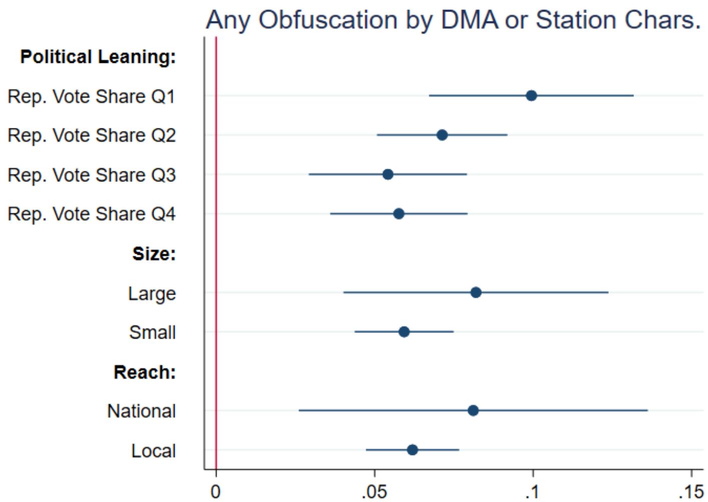
This figures plots the number of incidents (Panel a) and new stories (Panel b) on police and civilian killings in our main analysis sample. Our sample includes incidents and news stories where a suspect was identified for civilian killings. Data sources: GVA, MPV and News Exposure.

Figure C.2: Percentage of Stories with Any Obfuscation

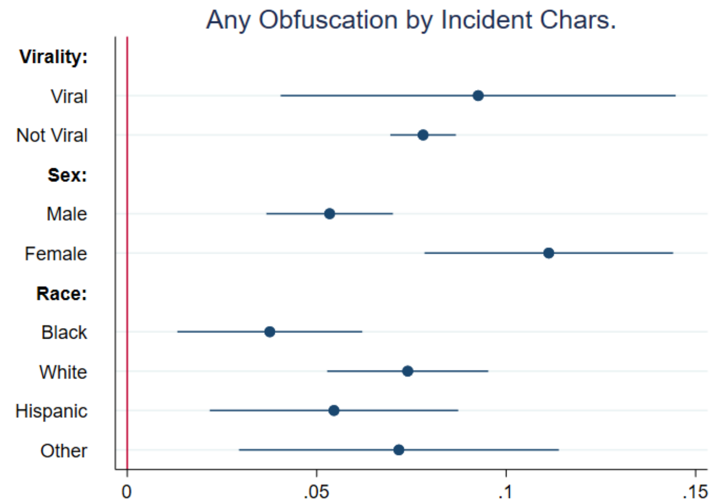


This figures presents the percent of stories over time that have some obfuscation, for police and civilian killings. Our sample includes incidents and news stories where a suspect was identified for civilian killings. We define obfuscation as having passive forms, no agent, intransitive verbs, or nominalization. See section 2 for more details. Source: News Exposure

Figure C.3: Heterogeneity in Obfuscation



(a) Heterogeneity by DMA or Station Chars.



(b) Heterogeneity by Incident Chars.

This figure presents differences in point estimates across different subgroups of news stories, by station characteristics (subfigure a) and victim characteristics (subfigure b). We estimate Equation 1 separately for each subgroup. The political leaning coefficients break out media markets by quartile of share republican vote in the DMA. Size depicts the media market size based on population, where 'Large' is the top 10 and 'Small' all the others. National versus local reach is based on the TV station having only local coverage or not. Virality is defined as having more than 100 stories per incident appearing in our sample. Lastly, the sex and race of defendants are based on information on victims from MPV and GVA. We define "obfuscation" as having passive forms, no agent, intransitive verbs, or nominalization. See section 2 for more details. Source: News Exposure, MPV, GVA

Appendix D Additional Tables

Table D.1: Descriptive Statistics Across Different Samples

	Police Killings			Civilian Killings		
	No Filter (1)	Subject Filter (2)	Story+Sent.Filter (3)	No Filter (4)	Subject Filter (5)	Story+Sent.Filter (6)
	Mean					
General Chars:						
Have Name of Victim	0.97	0.99	1.00	0.84	1.00	1.00
Age	36.80	36.77	36.79	32.15	33.31	33.90
Male	0.95	0.95	0.95	0.82	0.75	0.72
Black	0.17	0.18	0.18	0.37	0.31	0.28
Hispanic	0.18	0.17	0.16	0.11	0.10	0.09
White	0.55	0.58	0.58	0.29	0.42	0.47
Other/Unknown	0.10	0.08	0.07	0.24	0.16	0.15
Share Vote Rep. DMA	0.48	0.48	0.49	0.47	0.51	0.52
Caused by Gunshot	0.95	1.00	1.00	1.00	1.00	1.00
MPV Chars:						
Body Camera	0.07	0.07	0.07	.	.	.
Victim Not Fleeing	0.44	0.45	0.46	.	.	.
Unarmed/Unknown	0.18	0.15	0.15	.	.	.
GVA Chars:						
Have Name of Suspect	.	.	.	0.72	1.00	1.00
Suicide	.	.	.	0.02	0.00	0.00
Domestic Violence	.	.	.	0.09	0.18	0.21
Murder and Suicide	.	.	.	0.07	0.11	0.14
Gang-related	.	.	.	0.04	0.04	0.02
Near School	.	.	.	0.00	0.00	0.01
Home Invation	.	.	.	0.02	0.04	0.03
Number Victims in Incident	.	.	.	1.24	1.37	1.40
Observations	7663	6070	5759	49277	14011	7943

Source: MPV; News Exposure. *Notes:* This table presents descriptive statistics on across different filters. represents the overall data across the MPV and GVA, except for filtering GVA data so it does not include police killings. is the sample once we filter by characteristics of the incident. is our main final sample, which includes a filter for sentences where both a victim and an alleged perpetrator were identified, the story being predicted to be about crime with probability over 30%, and so on. Data sources: GVA, MPV and News Exposure.

Table D.2: Additional robustness tests

Panel A: Story Level	All Stories	
Dimension of Obfuscation:	Any (1)	No Explicit Agent (2)
Police Killing	0.085*** (0.009)	0.026*** (0.009)
Observations	192,944	192,944
Mean Civ.Shoot.	0.5366	0.2927

Panel B: Sentence Level	All Sentences		1st Sentence	
Dimension of Obfuscation:	Any (1)	No Explicit Agent (2)	Any (3)	No Explicit Agent (4)
	Panel B.1: Years of Overlap GVA-MPV (2014-2018)			
Police Killing	0.076*** (0.009)	0.031*** (0.005)	0.124*** (0.011)	0.068*** (0.007)
Observations	350,023	350,023	134,742	134,742
Mean Civ.Shoot.	0.2935	0.1354	0.2814	0.1266
	Panel B.2: Killings Appearing In Two Days or More			
Police Killing	0.081*** (0.008)	0.030*** (0.005)	0.127*** (0.010)	0.065*** (0.006)
Observations	443,369	443,369	172,698	172,698
Mean Civ.Shoot.	0.2875	0.1313	0.2757	0.1222
Controls	Story Controls+DMA FE			

Notes: Standard errors clustered by subject (*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$). This table presents additional robustness tests for our main results, for all sentences (columns 1 and 2) and for the first sentence (columns 3 and 4). Panel A presents results for the years in which we have both MPV and GVA data. Panel B limits our sample to killings that appear for at least two separate calendar days in the news cycle. We define 'obfuscation' as having passive forms, no agent, intransitive verbs, or nominalization. We define 'No Explicit Agent' as having no agent or intransitive verbs. See section 2 for more details. Source: News Exposure.

Table D.3: Online experiment: balance test.

	Age (1)	Male (2)	Prolific Experience (3)	Non-US Born (4)	Black (5)	White (6)	Hispanic (7)	Asian (8)	College Degree (9)	Passed Attention (10)
Passive	1.11	0.00	47.94	-0.01	-0.00	-0.01	-0.00	0.02	0.01	0.01
	(0.80)	(0.03)	(38.11)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)
No Agent + Nominalization	0.65	0.01	64.40*	-0.00	0.02	0.01	0.00	-0.00	-0.01	0.01**
	(0.80)	(0.03)	(37.98)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)
Intrasitive + Nominalization	0.92	-0.02	46.97	-0.01	-0.02	0.00	0.02	-0.00	-0.03	0.01
	(0.80)	(0.03)	(38.23)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)
Weapon	-0.86	0.00	17.67	0.00	-0.01	-0.03	-0.01	0.03**	-0.02	0.00
	(0.57)	(0.02)	(27.06)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.00)
Mean Dep. Var.	36.51	0.44	881.85	0.06	0.09	0.77	0.08	0.11	0.61	0.98
SD Dep. Var.	13.88	0.50	621.81	0.24	0.29	0.42	0.27	0.31	0.49	0.13
N	2397	2402	2397	2402	2402	2402	2402	2402	2402	2402

Notes: The outcomes in each column capture respondent characteristics. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.4: Online experiment: Additional results on the effect of the narrative structure on the judgment of the event

	Officer Justified (1)	Negative Sentiment (2)
	Panel A: No mention of victim weapon	
Passive	0.22	-0.07*
	(0.14)	(0.03)
No Agent + Nominalization	0.28**	-0.17***
	(0.14)	(0.04)
Intrasitive + Nominalization	0.36***	-0.17***
	(0.14)	(0.04)
Mean Dep. Var.	2.60	0.35
SD Dep. Var.	1.71	0.48
N	1201	1201
	Panel B: Victim has a weapon	
Passive	-0.14	-0.04
	(0.14)	(0.03)
No Agent + Nominalization	0.21	-0.11***
	(0.14)	(0.03)
Intrasitive + Nominalization	0.29**	-0.08***
	(0.14)	(0.03)
Mean Dep. Var.	3.49	0.21
SD Dep. Var.	1.79	0.41
N	1201	1201

Notes: Panel A, we presents results when we do not specify that the victim was armed, while we do in Panel B. The outcome in column 1 is agreement on a scale from 1 to 7 with the statement that the police officer was justified in shooting the person. The outcome in column 2 is a dummy equal to 1 if the person thought that the police officer was depicted negatively in the story. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.5: Online experiment: heterogeneity by race

	Moral Reponsibility		Department Penalty		Legal Penalty		Yearly Police Killings		Donation Reform		Support Reform	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Black	0.16**	0.15***	0.33	0.30	0.36	0.35*	77.46	-22.34	11.37***	6.44**	0.05	0.08**
	(0.07)	(0.05)	(0.29)	(0.21)	(0.30)	(0.21)	(158.93)	(112.63)	(4.33)	(3.07)	(0.05)	(0.03)
Obfuscation	-0.06**		-0.19*		-0.19*		-103.68*		-1.74		-0.01	
	(0.02)		(0.10)		(0.10)		(55.19)		(1.50)		(0.02)	
Interaction	-0.08		-0.00		0.11		152.68		-4.46		0.05	
	(0.08)		(0.34)		(0.34)		(184.09)		(5.01)		(0.06)	
No Explicit Agent		-0.07***		-0.28***		-0.32***		-171.78***		-2.82**		-0.01
		(0.02)		(0.09)		(0.09)		(47.98)		(1.31)		(0.01)
Interaction		-0.09		0.07		0.17		430.82***		3.25		0.01
		(0.07)		(0.29)		(0.30)		(160.01)		(4.37)		(0.05)
Mean Dep. Var.	0.87	0.87	4.24	4.24	4.05	4.05	1423.05	1423.05	77.82	77.82	0.91	0.91
SD Dep. Var.	0.34	0.34	2.42	2.42	2.50	2.50	1285.59	1285.59	23.51	23.51	0.29	0.29
N	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402

Notes: Obfuscation is equal to 1 if the sentence structure is passive, no agent, or intransitive. No explicit agent is equal to 1 if the sentence structure is no agent or intransitive. The outcome in columns 1-3 is a dummy equal to 1 if the respondents think that the police officer is morally responsible for the victim's death. The outcomes in the remaining columns are support, on a scale from 1 to 7, for department penalties (columns 4-6) and legal penalties (columns 7-9), respectively. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.6: Online experiment: heterogeneity by sex

	Moral Reponsibility		Department Penalty		Legal Penalty		Yearly Police Killings		Donation Reform		Support Reform	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female	0.13***	0.09***	0.34**	0.31***	0.33*	0.31**	278.68***	179.55***	4.40*	4.14**	0.06**	0.05**
	(0.04)	(0.03)	(0.17)	(0.12)	(0.17)	(0.12)	(91.35)	(64.90)	(2.50)	(1.77)	(0.03)	(0.02)
Obfuscation	-0.03		-0.17		-0.16		-5.11		-1.30		0.00	
	(0.03)		(0.15)		(0.15)		(79.05)		(2.16)		(0.02)	
Interaction	-0.06		-0.04		-0.04		-152.84		-1.53		-0.01	
	(0.04)		(0.19)		(0.20)		(105.83)		(2.89)		(0.03)	
No Explicit Agent		-0.08***		-0.28**		-0.30**		-118.16*		-1.56		-0.01
		(0.03)		(0.13)		(0.13)		(68.94)		(1.88)		(0.02)
Interaction		-0.00		0.00		-0.02		-28.47		-1.76		-0.00
		(0.04)		(0.17)		(0.17)		(92.20)		(2.52)		(0.03)
Mean Dep. Var.	0.78	0.78	4.09	4.09	3.87	3.87	1475.96	1475.96	69.41	69.41	0.89	0.89
SD Dep. Var.	0.41	0.41	2.16	2.16	2.19	2.19	1253.06	1253.06	29.43	29.43	0.31	0.31
N	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402

Notes: Obfuscation is equal to 1 if the sentence structure is passive, no agent, or intransitive. No explicit agent is equal to 1 if the sentence structure is no agent or intransitive. The outcome in columns 1-3 is a dummy equal to 1 if the respondents think that the police officer is morally responsible for the victim's death. The outcomes in the remaining columns are support, on a scale from 1 to 7, for department penalties (columns 4-6) and legal penalties (columns 7-9), respectively. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.7: Online experiment: heterogeneity by college

	Moral Reponsibility		Department Penalty		Legal Penalty		Yearly Police Killings		Donation Reform		Support Reform	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
College	0.01	0.02	0.11	0.07	0.07	0.04	7.99	-76.30	-2.85	-1.50	0.02	-0.01
	(0.04)	(0.03)	(0.17)	(0.12)	(0.17)	(0.12)	(93.26)	(66.47)	(2.54)	(1.81)	(0.03)	(0.02)
Obfuscation	-0.03		-0.12		-0.12		-43.13		-2.98		0.03	
	(0.03)		(0.15)		(0.16)		(84.19)		(2.30)		(0.03)	
Interaction	-0.05		-0.10		-0.10		-77.44		1.35		-0.05	
	(0.04)		(0.20)		(0.20)		(107.97)		(2.94)		(0.03)	
No Explicit Agent		-0.03		-0.23*		-0.25*		-162.95**		-2.11		-0.00
		(0.03)		(0.13)		(0.14)		(73.05)		(1.99)		(0.02)
Interaction		-0.09**		-0.07		-0.08		47.35		-0.77		-0.01
		(0.04)		(0.17)		(0.17)		(93.89)		(2.56)		(0.03)
Mean Dep. Var.	0.73	0.73	3.98	3.98	3.76	3.76	1355.68	1355.68	66.35	66.35	0.87	0.87
SD Dep. Var.	0.44	0.44	2.15	2.15	2.18	2.18	1183.28	1183.28	31.53	31.53	0.33	0.33
N	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402	2402

Notes: Obfuscation is equal to 1 if the sentence structure is passive, no agent, or intransitive. No explicit agent is equal to 1 if the sentence structure is no agent or intransitive. The outcome in columns 1-3 is a dummy equal to 1 if the respondents think that the police officer is morally responsible for the victim's death. The outcomes in the remaining columns are support, on a scale from 1 to 7, for department penalties (columns 4-6) and legal penalties (columns 7-9), respectively. See Appendix B for the full questions. We report the mean and standard deviation of the dependent variable for the Active narrative structure. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.8: Obfuscation in news of police killings, depending on whether there was body camera and if victim was allegedly fleeing

	All Sentences		1st Sentence	
Dimension of Obfuscation:	Any (1)	No Explicit Agent (2)	Any (3)	No Explicit Agent (4)
	Panel A: Body Camera			
Police Killing	0.116*** (0.016)	0.047*** (0.009)	0.154*** (0.018)	0.098*** (0.013)
Observations	175,993	175,993	68,163	68,163
Mean Civ.Shoot.	0.2935	0.1354	0.2814	0.1266
	Panel B: No Body Camera			
Police Killing	0.070*** (0.008)	0.025*** (0.005)	0.119*** (0.010)	0.058*** (0.006)
Observations	437,237	437,237	169,717	169,717
Mean Civ.Shoot.	0.2935	0.1354	0.2814	0.1266
Controls	Story+DMA FE			

Notes: Standard errors clustered by subject (*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$). This table breaks out differential obfuscation for stories about police killings, relative to stories about civilian killings, depending on whether there was a body camera in the police killing or if the victim was fleeing as defined by MPV. Panel A is for cases in which there was a body camera, while Panel B is for cases in which there was no body camera. Our sample includes incidents and news stories where a suspect was identified for civilian killings. We define ‘obfuscation‘ as having passive forms, no agent, intransitive verbs, or nominalization. We define ‘No Explicit Agent‘ as having no agent or intransitive verbs. See Section 2 for more details. Source: News Exposure.