Supplementary Materials: What Counts as Terrorism? Racial heuristics and media portrayals of mass shooters

Emily K. Gade,[1] Dallas Card,[2] Sarah K. Dreier,[3] and Noah A. Smith[4]

[1]Acting Assistant Professor or Political Science and Moore/Sloan Innovations in Data Science Postdoctoral Fellow, *University of Washington*

[2]Ph.C., Machine Learning Department, *Carnegie Mellon University*

[3]Postdoctoral Research Fellow, Department of Political Science, *University of Washington*

[4]Professor, Computer Science and Engineering, *University of Washington*

Contents

List of Figures

This document presents additional information about data collection and coding, descriptive statistics and data exploration, as well as the regression table for the models presented in the article. It proceeds as follows: Appendix A presents the regression tables for Models 1-7 in the article. Appendix B presents additional background data and data collection details regarding the Mass Shootings in America Database and our original news corpus. Appendix C provides supplementary material for the statistical analysis, including a complete coding ontology for all variables used and descriptive statistics. Appendix D provides supplementary material for the qualitative analysis, including a description of each article coded and the coding schema used. Finally, Appendix E includes details about the Scholar model and our unsupervised analysis.

A Regression tables for Models 1-7 (Figures 5/6) in article

Table 1: OLS Models of a count of terrorism mentions and "racalized perpetrator"

	Dependent variable:			
	# Terror Mentions			
	(Model 1)	(Model 2)	(Model 3)	(Model 4)
# of Victims	0.535***	0.443***	0.432***	0.534***
	(0.096)	(0.086)	(0.086)	(0.095)
# of Articles	0.084***	0.075***	0.077***	0.085***
	(0.019)	(0.017)	(0.017)	(0.019)
Group Affil/Contact	16.970***	11.310**	11.305**	17.191***
	(5.352)	(4.863)	(4.850)	(5.335)
Gov or Political Target	−0.256	−2.350	−2.125	0.324
	(2.606)	(2.351)	(2.342)	(2.594)
Domestic/Social Dispute	0.741	−0.015	0.579	−0.250
	(1.274)	(1.137)	(1.131)	(1.306)
Target is a School	−2.730*	−2.756**	−1.989	−2.517*
	(1.516)	(1.357)	(1.356)	(1.510)
Target Any Minority	3.508	5.656**	5.605**	4.704*
	(2.807)	(2.518)	(2.511)	(2.820)
Event is Terrorism	3.062*	1.657	1.802	2.790*
	(1.634)	(1.478)	(1.472)	(1.630)
Immigrant Shooter	1.406			
	(2.006)			
Middle Eastern Shooter		29.047***		
		(3.707)		
Muslim Shooter			29.303***	
			(3.693)	
White Shooter				−3.466***
				(1.232)
Constant	−4.617***	−3.722***	−4.096***	−2.690**
	(1.048)	(0.934)	(0.928)	(1.223)
Observations	265	265	265	265
R^2	0.570	0.652	0.654	0.582
Adjusted R^2	0.554	0.640	0.642	0.567
Residual Std. Error	9.461	8.501	8.481	9.430
	(df = 255)	(df = 255)	(df = 255)	(df = 255)
F Statistic	37.494***	53.191***	53.585***	38.564***
	(df = 9; 255)	(df = 9; 255)	(df = 9; 255)	(df = 9; 255)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2: OLS models of whether the news and expert coding disagree (5 is high disagreement, 0 is perfect agreement)

	Dependent variable:		
	Mis-Categorized		
	(Model 5)	(Model 6)	(Model 7)
Number of Victims	0.002	0.004	0.005
	(0.012)	(0.012)	(0.012)
Group Affil/Contact	0.185	0.298	0.408
	(0.612)	(0.624)	(0.622)
Gov or Political Target	0.816***	0.867***	0.919***
	(0.294)	(0.296)	(0.296)
Domestic/Social Dispute	−0.514***	−0.520***	−0.504***
	(0.158)	(0.156)	(0.156)
Target was a School	−0.301	−0.311*	−0.301
	(0.186)	(0.186)	(0.185)
Target was Any Minority	0.734**	0.704**	0.684**
	(0.322)	(0.321)	(0.319)
# Valid Articles	0.003	0.003	0.004
	(0.003)	(0.003)	(0.003)
Immigrant Shooter	0.049		
	(0.250)		
Muslim Shooter		−0.439	
		(0.510)	
Middle Eastern Shooter			−0.854*
			(0.507)
Constant	0.650***	0.651***	0.638***
	(0.130)	(0.128)	(0.128)
Observations	265	265	265
R^2	0.157	0.159	0.166
Adjusted R^2	0.131	0.133	0.140
Residual Std. Error (df = 256)	1.179	1.178	1.173
F Statistic (df = 8; 256)	5.959***	6.063***	6.373***
Note:	*p<0.1; **p<0.05; ***p<0.01		

B Case Background: Mass Shootings in America Database and Developing Our News Corpus

B.1 News Article Filtering Mechanisms

Data collection of newspaper articles related to shootings in the MSA database involved the following process. Following the contention established by (Card et al., 2015)'s media frames project, we examined articles among the following eight mainstream national or regional news sources: *San Jose Mercury News, USA Today, St. Louis Post-Dispatch, Daily News, Tampa Bay Times, Washington Post, Philadelphia Inquirer, Saint Paul Pioneer Press, Palm Beach Post, Atlanta Journal and Constitution, New York Times,* and *St. Petersburg Times*.

To be included, an article must have 1) been published within one week of the event and 2) received Lexis-Nexis tags with the state in which the event occurred, 3) receive Lexis-Nexis tags with at least one of the following: ("Mass shootings," "Shootings," "Terrorism," or "Terrorist attacks"), 4) explicitly mentioned the city in which the event occurred. Although likely omitting some relevant articles (false negatives) and including some unrelated coverage (false positives), we believe this filtering method provides a reliable corpus of news coverage.

B.2 Shift in usage of "Shooting" over time?

As part of our interrogation of relevant terms over time (see Section 1.1 "Terrorism's social meaning over time" in the article), we investigated whether the term "shooting" had experienced any shift in frequency of use in the New York Times corpus. If we plot the usage of the term "shooting", we find very little change in average frequency over this time period (Figure 1). Note that for this analysis, we don't include articles which include the term "game" or "film", as these drive large seasonal variations.

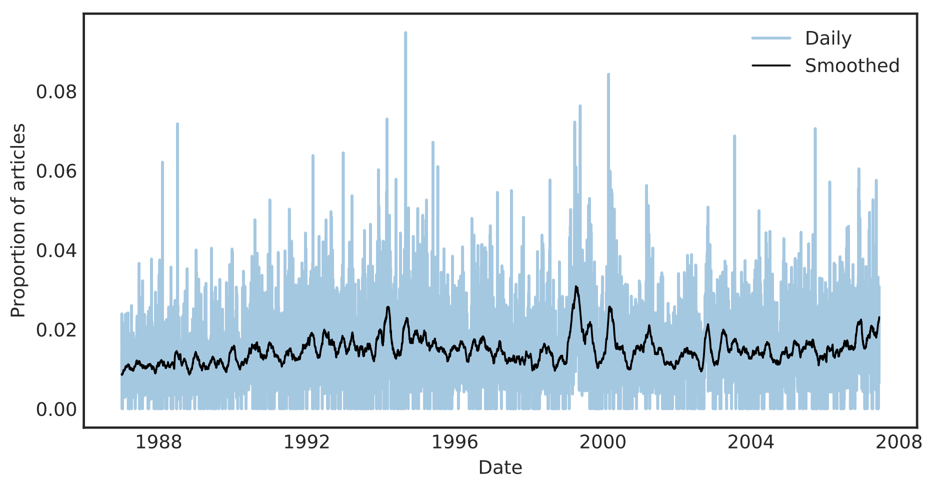

Figure 1: Proportion of *New York Times* articles in the NYT annotated corpus that contain the word "shooting", but not "game" or "film".

B.3 Discussion of Shootings in the Mass Shootings in America Database

Yearly incident counts have increased considerably since 1990 (Figure 2).[5] The majority of shootings involved ten deaths or fewer (Figure 3). Our analyses exclude the 35 cases lacking an identified perpetrator as well as 32 cases which occurred before 1990 and our available news data (n=265).

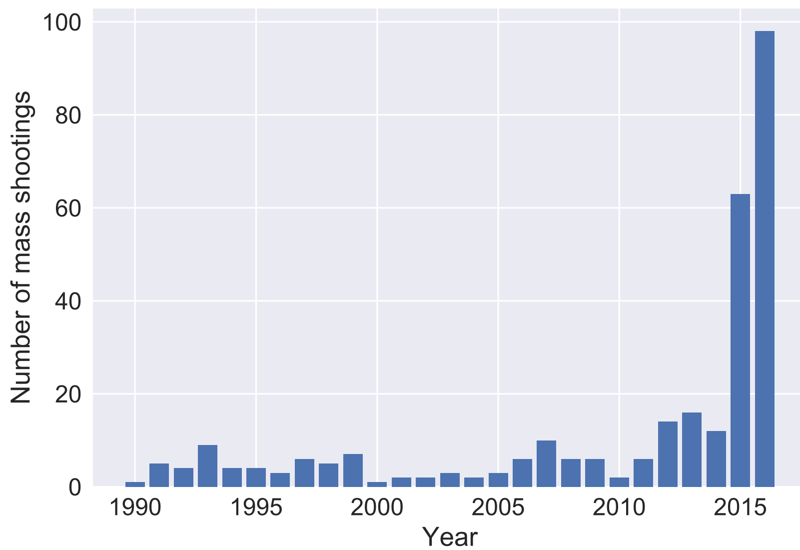

Figure 2: Number of events in the Stanford Mass Shootings database per year.

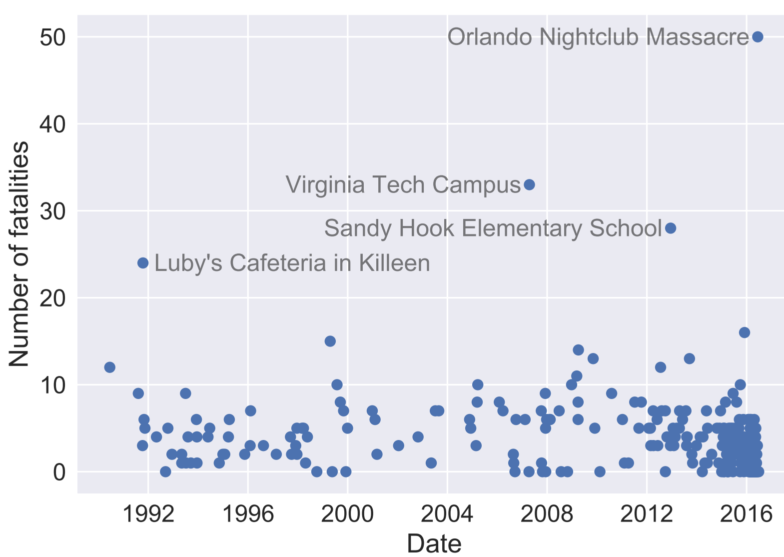

Figure 3: Number of fatalities for each incident.

[5]This may reflect increases in reporting and national awareness over time (i.e., reporting bias). However, the Federal Bureau of Investigation reports similarly dramatic increases in recent years (Grow, 2014).

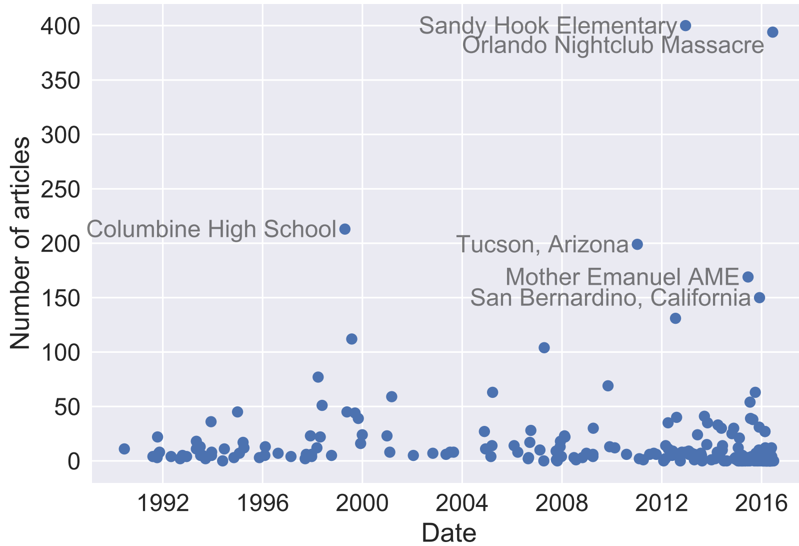

Figure 4: Number of articles about each event in the Stanford Mass Shootings database.

News coverage is notably skewed toward a few high-profile events, each of which generated hundreds of articles while others received no national coverage (Figure 4). Two events—the 2016 Pulse Nightclub shooting and the 2012 Sandy Hook shooting—were each covered in approximately 400 articles in the week following the event. In each case, this heightened coverage comports with research expectations. The Pulse shooting was the most deadly incident in our dataset (which excludes the 2017 shooting in Las Vegas that killed 58 people because these data end in 2016) and was perpetrated by a racialized minority who expressed an (unverified) affiliation with ISIS. The Sandy Hook perpetrator conducted an almost inconceivable attack on six and seven-year-old children, twenty of whom died, with no expressed motive. Beyond these exceptional events, there is not a strictly determined, consistent relationship between severity (e.g., number of casualties) and news coverage (Figure 3).

C Supplementary Material: Statistical Analysis

C.1 Coding Instances of Terrorism

We used an expert coder to qualitatively code each mass shooting in our database and evaluate whether or not the event could be considered terrorism under the broadest possible definition. We created a second variable which was graded (2, the event is definitely terrorism; 1 the event could be considered terrorism; 0 the event is definitely not terrorism). Our coding criteria for these designations focused on whether: (1) the event was violent (all were), and (2) whether or not the event could be considered to have a political motive beyond the event itself.

Our expert coder begin with the news stories associated with the incident in the Mass Shootings database. She then conducted a google search about the incident. Where possible, she used documents from court cases or other non-news sources (e.g. `https://schoolshooters.info/search-database`). She recorded every website, document or news source she looked at that was not one of the Mass Shootings in America new stories. Section F of this document contains the complete coding ontology used for all new variables.

Finally, we evaluate whether an event was "mis-classified"—whether an event could be considered terrorism under an expansive definition *and* it was called terrorism in the news, or whether the two measures disagree. An event was coded as 0, perfectly agree, if both the news and the expert coding agreed that an event was not terrorism (no event had 100% of news stories consider an event terrorism). An event was coded as "strongly agree" (1) if an event was considered terrorism under a broad definition and more than 35% of newspaper article talked about it that way, OR an event was *not* considered terrorism under a broad definition and fewer than 5% of article talked about it that way. An event was coded as "mildly agree" (2) if an event was terrorism and between 20-35% of news stories talked about it that way OR an event was not considered terrorism and between 5-10% of stories talked about it that way. An event was coded as "mild disagree" (3) if an event was considered terrorism and between 10-20% of stories talked about it that way, or an event was not considered terrorism and between 10-20% of stories talked about it that way. An event was coded as "strongly disagree" (4) if an event was considered terrorism and less than 10% of the news stories talked about it that way, or an event was not terrorism and more than 20% of news stories talked about it that way. Finally, an event was coded as "perfectly disagree" if an event could be considered terrorism under a broad definition, but no new stories mention it that way. There were no cases of the reverse (an event is not terrorism and all news stories mention it that way). See Table 3 for detailed description of coding of other variables.

C.2 Coding Ontology, All Variables, Statistical Analysis

The statistical analysis excludes 34 incidents in which the shooter was never identified (as makes it impossible to code demographic/motive data). It also excluded the 32 mass shooting incidents which occurred before 1990 (as we were not able to collect machine-readable news data for those incidents). There were several overlapping events in the MSA database: 250/282, deleted event 282; 253/286, 286 deleted; 260/303, 303 deleted; 255/290, 290 deleted; 256/293, 293 deleted. We created 37 new variables in addition to the Mass Shooting in America Database variables. These variables, and the ontology we used to code each, is described below:

Variable Name	Description	Examples

Shooter Race	Includes category for middle eastern, validates MSA codes, states if coded from photo	
Middle Eastern Shooter	Binary, based on country of origin	
White Shooter	Excludes middle eastern, Hispanic	
Shooter Immigrant	Binary, explicated mention of immigrant status = 1	
Motive Jihad	Binary, explicit mention of Jihad motive =1	
Shooter Interpreted as Muslim	binary, did the news interpret the shooter as Muslim= 1 (one case of mis-identification that was later corrected)	
Target Muslim/Shik	Binary, target was Muslim/shik? = 1	
Target: Racial Minority	Binary, targeted because they were a racial minority = 1	
Target: Religious Minority	Binary, targeted because they were a religious minority = 1	
Target: Women	Binary, targeted because they were women = 1	
Target: Immigrant	Binary, targeted because they were an immigrant = 1	
Target: LGBTQ	Binary, targeted because they were LGBTQ = 1	
Target: Any Minority	Binary, targeted fit in any of the previous five columns = 1	
Target: Gov	Binary, target was a member of the gov/part of gov = 1	includes TSA, congress, post office, police
Target: Religious Group	Binary, target was a place of faith (any faith) = 1	
Target School	Binary, target was a school =1	Includes graduations, classrooms, school sporting events
Right Wing	Binary, perp was right wing = 1	perp espoused extreme right-wing views (e.g. KKK/white supremist)

Left Wing	Binary, perp was left wing = 1	perp known to espouse extreme left wing views (e.g. environmental terrorism)
Group Affiliation/Contact	Binary, perp was affiliated with a political group that has at any point used violence = 1	Includes kkk, neo nazi groups, ISIS and black panthers
Motive Economic	Binary, perp had an economic motive = 1	Includes debt, bankruptcy, feeling of unfair pay
Motive Copy Cat	Binary, perp stated another incident was a motive for this one = 1	Most common – copying columbine
Perceived Grievance, Domestic/Social Dispute	Binary, perp was involved in social dispute/perpetrating domestic violence = 1	
Perceived Grievance, School/Bullying	Binary, perp felt they had been bullied or had a school related grievance = 1	Also includes failing grades, frustration with degree, bullying or exclusion
Perceived Grievance, Work	Binary, perp had a work-related grievance = 1	
Perceived Grievance, Racism/Discrimination	Binary, perp felt they were being discriminated against = 1	Includes general frustration with racism
Perceived Grievance, All	Binary, perp described as having any perceived grievance = 1	Includes domestic, social, school, work, or discrimination related grievance
Motive, Hate Crime	Binary, perp was perpetrating a hate crime = 1	
Motive, Gov/Political Target	Binary, perp was targeting the government for a political aim =1	
Motive, Fame	Binary, perp was seeking fame =1	Needed explicit mention of this from suicide note or perp statement
Motive, Suicide	Binary, perp was trying to commit suicide = 1	Needed suicide note or some evidence that the perp was trying to commit suicide before the attack
Motive, Environmental	Binary, perp had an environmental motive (no instances)	

Motive, Social Issue	Binary, perp had a social motive = 1	Includes anti-woman, anti-abortion, racism, wanting to start a race war/pro-racism, etc.
Gang?	Binary, was this incident described as having gang involvement? = 1	There aren't supposed to be any of these in the MSA, but we found a few
Motive, Drug/Alcohol	Binary, perp was high or drunk during the incident = 1	
Motive, Mental Illness	Binary, perp had a documented mental illness and that was described as being part of the cause of the incident = 1	
Accident?	Binary, this incident was understood in the news as an accident = 1	
Code as Terrorism?	Coded from 0-2, did this incident definitely meet the concatenated definition of terrorism? = 2; Might be terrorism? = 1; Definitely not terrorism = 0	Had to include a defined political motive (e.g. manifesto, stated political goal)
Terrorist Group Claims Incident?	Binary, did a terrorist group claim the incident or support the incident? = 1	Almost no cases of this
Mental Illness History?	Binary, did the perp have a history of mental health problems? = 1	

Table 3: Coding Ontology for Statistical Analysis, Mass Shootings Events

C.3 Descriptive Statistics of Variables in Statistical Analysis

Table 4: Descriptive Statistics of Key Variables

Variable	N = 265
Event is Terrorism	
min/max	0, 2
mean (sd)	0.16 ± 0.52
Count of Terrorism Mentions	
min/max	0, 203
mean (sd)	2.08 ± 14.17
Disagreement/Miscategorization	
min/max	0, 5
mean (sd)	0.58 ± 1.26
Prep Race/Demographic	
Middle Eastern Shooter*	7
Muslim Shooter*	7
Immigrant Shooter	27
White Shooter	120
Target Type	
Gov. Target	19
Domestic/Social Dispute	95
School Shooting	55
Minority Target	17

* *Note: there is significant but imperfect overlap between these two categories.*

D Supplementary Material: Qualitative Analysis

D.1 Article Selection

We conducted a qualitative comparative analysis of Nidal Hasan's mass shooting at Fort Hood (2009) and Dylann Roof's mass shooting at Mother Emanuel A.M.E. Church (2015). We collected all articles published in the *New York Times* about each attack in roughly the week following the incident (8-9 days). Given our focus on perpetrator coverage, articles were identified by searching for the perpetrator's first and last name. We included news reports, news blogs, daily/weekly news reviews, opinion pieces, letters to the editor, expert debates, *Times* magazine pieces, and educational features. We excluded from our analysis photos, videos, podcasts, interactive timelines, social media links, Spanish language coverage, *International New York Times* articles (which largely duplicated domestic coverage), and reader comments. These criteria yielded 50 articles covering Hasan's attack and 49 covering Roof's attack.

Using Atlas.ti, we read and hand-coded all articles published within the first two days of each attack (Hasan: Nov 5-6, 2009; Roof: June 18-19, 2015[6]) and all articles in the first week that mentioned "terrorism," "terrorist," or "terror." These criteria yielded 30 coded articles covering Hasan's attack and 29 covering Roof's attack.

D.2 Coding Ontology: Racial Heuristics

We coded coverage of Hasan and Roof to compare the extent to which media racialized each perpetrator. In theory, an article that makes no reference to racial identity (e.g., "Muslim," "white man") would be one that evokes no racial heuristics. Our results yielded more than 100 instances in which coverage racialized Hasan's identity (as a Muslim, a Middle-Eastern American, a son of immigrants, or having family members living in Palestine) and only twenty-two instances in which Roof was racialized (as a white man or someone with a Lutheran Christian background). We therefore conclude that, relative to Roof, Hasan's racial identity played a central role in how media covered his attack.

This coding ontology sought to distinguished between references to perpetrators' **racialized identities** (which we considered *irrelevant* to perpetrator motives) and any **extreme political ideology** that may have motivated their violence (which we considered *relevant* to perpetrator motives). In many cases, this distinction was relatively straightforward:

- "Major Hasan prayed every day at the mosque [RACIALIZED IDENTITY]."
- Hasan's "parents, Palestinians who had immigrated from the West Bank in the 1960s [RACIALIZED IDENTITY], moved the family to Roanoke when he was a youth."
- Hasan's "dozen or so messages to the cleric, Anwar al-Awlaki [EXTREME POLITICAL IDEOLOGY], were largely questions about Islam [RACIALIZED IDENTITY], not expressions of militancy."
- "The massacre of nine African-Americans in Charleston has been classified as a possible hate crime [MOTIVE: HATE], apparently carried out by a 21-year-old white man [RACIALIZED IDENTITY] who once wore an apartheid badge and other symbols of

[6] Roof's shooting was conducted in the evening of June 17, 2015, and he was not identified until the next day. Therefore, coverage of the attack with Roof as the perpetrator did not begin until June 18.

white supremacy [EXTREME POLITICAL IDEOLOGY]. But many civil rights advocates are asking why the attack has not officially been called terrorism [TERRORISM MENTIONED]."

In many other cases, however, these distinctions were difficult to objectively identify. Do questions about whether Hasan was a '*jihadist*' appropriately capture the relevant contours of his attack, or do they represent implicit heuristic assumptions that erroneously connect his Muslim faith to his actions (which experts acknowledge had multiple motivations)? Do commentary that situate Roof's attack within an American history that institutionalized and mainstreamed white supremacist ideologies appropriately represent Roof's motivations or inappropriately implicate white-identifying people as terrorists? Amid this uncertainty, we attempted to code contextual references that appeared to unreasonably assume or connect the perpetrator's race or religion to his (or other) extreme ideologies or violent attacks as "racial identity." We similarly attempted to code contextual references that appeared to be relevant references to extreme connections or motivations as "extreme political ideology."

In making these coding distinctions, we considered it relevant that Roof explicitly named his extreme political motivations before his attack and that his online white-supremacist presence was unambiguous. Therefore, there was no uncertainty in his motivations. In contrast, Hasan's motives were multifaceted and not clearly articulated. In the absence of this evidence (particularly initially, before more information was known), references that presumptuously associated Islam, terrorism, and Hasan's attack were coded as "racial identity" and treated as an indication that the perpetrator was racialized. Table 5 provides code frequencies.

D.3 Coding Ontology: Terrorism Mentions

Next, we compare how media associated Hasan and Roof with terrorism. Under our operational definition, we consider both Hasan's and Roof's attacks as terrorism. In both cases, officials investigated whether the event should be considered terrorism but announced no definitive conclusion. In neither case were perpetrators tried in court for terrorism. We therefore would expect media mentions of terrorism to be relatively comparable among the two cases, provided that racial or other subjective treatments were not shaping these results. Indeed, the percentage of Hasan articles referencing terrorism (18 articles, or 36%) is only slightly higher than the percentage of Roof articles referencing terrorism (15 articles, or 30.6%).

However, this does not demonstrate how terrorism was discussed, or the centrality that terrorism frames played, in the coverage of these articles. We therefore coded and compared the extent to which media associated Hasan's and Roof's attacks with terrorism. We distinguished four sub-categories of terrorism references (Table 5):

- **Terrorism mentioned**: "Can we call his attack an act of terrorism?" or "[The case] is being investigated by the Justice Department as a possible case of domestic terrorism."
- **Terrorism implied**: "The system set up after Sept. 11, 2001,... failed to stop the deadly episode." or "Mr. Awlaki was... preaching at the Dar al-Hijrah Islamic Center... [T]hree of the [9/11] hijackers attended service there."
- **Terrorism asserted**: "This was an act of racial terrorism."
- **Terrorism negated**: "Investigators have tentatively concluded that he... was not part of a terrorist plot."

	Hasan	Roof
Total articles	50	49
Articles referencing terrorism (keyword search)	18 (36%)	15 (30.6%)
Articles referencing perpetrator religion (keyword search)	26 (52%)	11 (22.4%)
Total Codes	330	306
Mentions of racialized identity (hand-coded)	106	22
Mentions of terrorism (hand-coded)	61	41
Terrorism mentioned	16	11
Terrorism implied	29	6
Terrorism asserted	2	24
Terrorism negated	14	0

Table 5: Coding Results for Comparative Qualitative Analysis

This coding approach allows us to capture both explicit references to terrorism and implicit references built into the tangential facts or incidents featured in the articles. The resulting analysis suggests that the frequency, centrality, presented objectivity, and commentary associated with these designations are considerably different among these two cases. Most notably, we coded 61 mentions of terrorism in reference to Hasan and only 41 mentions of terrorism in reference to Roof.

Furthermore, many references in Hasan's coverage reported on law-enforcement investigations into whether his actions were driven by a "radicalization" of his faith and by communications with known "Islamic" terrorist entities. Coverage of Roof's attack made far fewer references to objective law-enforcement investigations about official terrorist designations. Instead, those treating Roof as a terrorist were disproportionately authors of editorial opinion pieces ("A Millennial Race Terrorist"), many of whom critiqued a disproportionate lack of official and media considerations of Roof's terrorism.

The disproportionately frequent references to official considerations about terrorism—and even the negation of those determinations—in Hasan's case (relative to Roof's) can serve as cues that reflect and reinforce racialized associations between Islam and terrorism. Even a higher frequency of negated references to terrorism attached to Muslim perpetrators (like Hasan) can reinforce these associations. We therefore conclude that, relative to Roof, Hasan's attack was more directly associated with terrorism on a number of dimensions.

D.4 *New York Times* Articles Quoted in Main Analysis

Article	**Quote**
Army Doctor Held in Ft. Hood Rampage (Nov 5, 2009)	"Military records indicated that Major Hasan was single, had been born in Virginia, had never served abroad and listed 'no religious preference' on his personnel records... The Muslim Public Affairs Council, speaking for many American Muslims, condemned the shootings as a 'heinous incident' and said, 'We share the sentiment of our president.' "
Updates on the Shootings at Fort Hood (Nov 6, 2009)	"The Council on American-Islamic Relations, or CAIR, a Muslim civil rights and advocacy group, issued this statement on Thursday night: 'We condemn this cowardly attack in the strongest terms possible and ask that the perpetrators be punished to the full extent of the law. No religious or political ideology could ever justify or excuse such wanton and indiscriminate violence... Along with innumerable condemnations of terror, CAIR has in the past launched an online anti-terror petition drive called "Not in the Name of Islam," initiated a television public service announcement (PSA) campaign against religious extremism and coordinated a "fatwa," or Islamic religious ruling, against terrorism and extremism.' "
Complications Grow for Muslims Serving in U.S. Military (Nov 8, 2009)	"Muslim leaders, advocates and military service members have taken pains to denounce the shooting and distance themselves from Major Hasan. They make the point that his violence is no more representative of them than it is of other groups to which he belongs, including Army psychiatrists. 'I don't understand why the Muslim-American community has to take responsibility for him,' said Ingrid Mattson, the president of the Islamic Society of North America. 'The Army has had at least as much time and opportunity to form and shape this person as the Muslim community.' "

The Rush to Therapy (Nov 9, 2009; opinion by David Brooks)	“That narrative has emerged on the fringes of the Muslim world. It is a narrative that sees human history as a war between Islam on the one side and Christianity and Judaism on the other. This narrative causes its adherents to shrink their circle of concern. They don’t see others as fully human. They come to believe others can be blamelessly murdered and that, in fact, it is admirable to do so... The conversation in the first few days after the massacre was well intentioned, but it suggested a willful flight from reality. It ignored the fact that the war narrative of the struggle against Islam is the central feature of American foreign policy. It ignored the fact that this narrative can be embraced by a self-radicalizing individual in the U.S. as much as by groups in Tehran, Gaza or Kandahar.”
Many Ask, Why Not Call Church Shooting Terrorism? (June 18, 2015)	“The massacre of nine African-Americans in Charleston has been classified as a possible hate crime, apparently carried out by a 21-year-old white man who once wore an apartheid badge and other symbols of white supremacy. But many civil rights advocates are asking why the attack has not officially been called terrorism. Against the backdrop of rising worries about violent Muslim extremism in the United States, advocates see hypocrisy in the way the attack and the man under arrest in the shooting have been described by law enforcement officials and the news media.”
On Racial Violence: Before Charleston’s Church Shooting, a Long History of Attacks (June 18, 2015; magazine piece by Douglas R. Egerton)	“In the coming days, the world will find out more about Dylann Storm Roof and his state of mind. But to dismiss him as simply a troubled young man is to disregard history. For 198 years, angry whites have attacked Emanuel A.M.E. and its congregation, and when its leaders have fused faith with political activism, white vigilantes have used terror to silence its ministers and mute its message of progress and hope. Denmark Vesey’s story should never be forgotten — nor should the tragedy of June 17.”
Returning Home to Console, Lindsey Graham Joins the Mourning (June 19, 2015)	“But on Friday morning, [Sen. Lindsey Graham] clarified when asked, saying he believed it was a racially motivated hate crime. ‘The only reason these people are dead is because they’re black,’ he said. Later, Mr. Graham described Mr. Roof as ‘a racial jihadist.’ ”

In Charleston, a Millennial Race Terrorist (June 21, 2015; opinion by Charles M. Blow)	"On Fox News's 'Fox and Friends,'... [one anchor said]: 'Extraordinarily, they called it a hate crime. Uh, and some look at it as, well, it's because it was a white guy, apparently, and a black church, but you made a great point just a moment ago about the hostility towards Christians.' Then there is the question of whether to call this terrorism. Terrorism, as commonly defined, suggests that the act must have some political motivation. (By defining it this way, we conveniently exclude that long legacy of racial terrorism as a political tool of intimidation and control in this country.) And yet, this case may even reach that bar."

E Supplementary Material: Scholar Framework

In order to perform unsupervised analysis of the character archetypes represented in coverage of mass shooters, while accounting for the skew in amount of coverage, and association with race, we make use of a framework for unsupervised modeling of documents with metadata, called SCHOLAR (Card, Tan and Smith, 2018).

As in topic models, this framework will allow us to learn a set of "topics" (clusters of semantically related words), and simultaneously infer a latent representation of each document in terms of these topics. As in Eisenstein, Ahmed and Xing (2011), topics are represented as per-word deviations from a background frequency.

In our application, the probability of a context word (an adjective referring to the perpetrator) occurring in an article is given by

$$p(w \mid \boldsymbol{\theta}_i) \propto \exp(\mathbf{b} + \boldsymbol{\theta}_i^\top \mathbf{B}^{(topic)} + \mathbf{c}_i^\top \mathbf{B}^{(covariate)}), \tag{1}$$

where $\boldsymbol{\theta}_i$ is a representation of the article on the k-dimensional simplex, $\mathbf{b}$ is a V-dimensional background term (where V is the size of the vocabulary), and $\mathbf{c}_i$ is a vector of one-hot covariates, indicating if the article is about one of the most frequent events. $\mathbf{B}^{(topic)}$ and $\mathbf{B}^{(covariate)}$ are weight matrices corresponding to positive and negative deviations from the background $\mathbf{b}$.

In addition, we can also optionally include a text classification component as an additional term in the objective function, such that we learn topics that are useful for predicting labels, in our case, race. This classification component takes the form of a neural network operating on the latent representation and covariates, i.e.,

$$p(y_i \mid \boldsymbol{\theta}_i) = f_y(\boldsymbol{\theta}_i, \boldsymbol{c}_i), \tag{2}$$

where y_i is the label for document i, and f_y represents a multi-layer perceptron.

Note that equation (3) has a similar form to the the structural topic model (Roberts et al., 2014), but SCHOLAR provides more scalable inference. It also allows us to incorporate pre-trained word vectors, which we do in this work, in order to obtain greater coherence in each dimension. Inference is performed using a variational autoencoder approach to Bayesian inference (please see Kingma and Welling (2014) and Card, Tan and Smith (2018) for details). As in topic models, the end result is a set of interpretable latent dimensions, each of which corresponds to a high and low probability words.

References

Card, Dallas, Amber E. Boydstun, Justin H. Gross, Philip Resnik and Noah A. Smith. 2015. The Media Frames Corpus: Annotations of Frames Across Issues. In *Proceedings of ACL*.

Card, Dallas, Chenhao Tan and Noah A. Smith. 2018. Neural Models for Documents with Metadata. In *Proceedings of ACL*.

Eisenstein, Jacob, Amr Ahmed and Eric P. Xing. 2011. Sparse Additive Generative Models of Text. In *Proceedings of ICML*.

Grow, Kory. 2014. "Mass Shootings Dramatically on the Rise, FBI Says.".
URL: *https://www.rollingstone.com/culture/culture-news/mass-shootings-dramatically-on-the-rise-fbi-says-67878/*

Kingma, Diederik P and Max Welling. 2014. Auto-Encoding Variational Bayes. In *Proceedings of ICLR.*

Roberts, Molly, Brandon Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Gadarian, Bethany Albertson and David Rand. 2014. "Structural Topic Models for Open Ended Survey Responses." *American Journal of Political Science* 58:1064–1082.