Demographic Correlates of Humanizing Language in Media Coverage of Crime: Evidence from the *Boston Globe*, 1976-84

Emily Ocasio[a,1] **and Tristan Potter**[b]

[a]The New School of Northern Virginia; [b]Drexel University

This manuscript was compiled on November 1, 2022

Media coverage of violent crime both reflects and shapes how society perceives members of different demographic groups. Measuring how media coverage varies across these groups—particularly those whose humanity society has historically failed to fully recognize—can therefore provide insight into public perceptions of different groups' humanity and the influence of the media in perpetuating such views. We merge FBI data on homicides in Massachusetts between 1976 and 1984 with the corresponding *Boston Globe* articles covering those homicides to study how the race, age, and sex of homicide victims and offenders interact to shape media coverage—both the quantity of coverage (number of articles written) and the quality of that coverage (use of humanizing language, determined through content analysis assisted by automated natural language processing). Our analysis, which controls for a battery of homicide-specific factors, reveals a rich pattern of differential media coverage across groups and subgroups. Specifically, we document that: (i) among male homicide victims, the probability of humanizing coverage is significantly higher for whites, with a difference of 30 percentage points for juveniles (<18) and a difference of between 10 and 15 percentage points for all other age groups; (ii) there is no statistically significant difference in the quantity or quality of coverage across races among female victims, except those between the ages of 18 and 29, among whom blacks are significantly less likely to receive humanizing coverage than whites; (iii) both female and juvenile victims tend to receive more coverage on average, and this coverage tends to be more humanizing when it occurs; and (iv) demographic characteristics of the offender do not appear to significantly influence coverage after controlling for other homicide-specific factors such as weapon, circumstances, etc. Our analysis highlights the importance of viewing differential media coverage of violent crime through an intersectional lens and how content analysis can be applied to produce comprehensive measures of humanizing language.

Media distortion analysis | Natural language processing | Intersectional analysis | Ideal victim theory

Crime stories make up 10-50% of all newspaper articles (1), with disproportionate attention devoted to stories about homicide in particular despite its relatively low prevalence (2). Among homicides, the quality and quantity of coverage varies substantially, and this variation in coverage is likely to relate systematically to the characteristics of the individuals involved in nuanced ways that have important social implications (1, 3, 4). Depiction of crime in newspapers can provoke public anxieties and direct those anxieties towards certain groups (5, 6), change jurors' decisions and trial outcomes (7), and affect attitudes of law enforcement officials towards both victims and perpetrators of crime (8, 9). Furthermore, in addition to influencing public perceptions and treatment of different groups, discrepancies in coverage of homicide also reflect the complex social forces from which they emanate.

Societal attitudes towards victims are affected by whether a victim fits the "ideal victim" prototype (weak, vulnerable, etc.) (10). Empirical evidence has shown that, when describing ideal victims, more sympathetic language is used (11, 12) and, in the case of newspaper reports of homicide, there is an increased level of coverage and sympathy (13). Principal among the characteristics influencing whether a victim is perceived as an ideal victim are demographic factors such as the victim's age, race, and sex (14) as well as the *interactions* between such factors. It is therefore important to analyze ideal victim typifications through an intersectional lens (12, 15). Failing to do so risks flattening both the lived experience of certain groups and our understanding of the complexity of the operative social forces. Accordingly, this paper aims to paint a nuanced picture of how demographic factors interact to shape the qualitative and quantitative coverage afforded to victims of violent crime.

We are not the first to study differential media coverage across demographic or socioeconomic groups. An important existing literature has sought to analyze media distortions by measuring, e.g., differences in news coverage of crime based on characteristics of victims and offenders (16–21). The studies that comprise this literature can be understood as varying principally along three key dimensions: (i) How media coverage is measured, (ii) which demographic groups are considered in the analysis, and (iii) data availability, including what other homicide-specific factors the analysis is able to control for, which homicides are chosen for analysis (and how they are matched with media coverage), and the geographic scope of the analysis.

Coverage is traditionally measured through readily available quantitative metrics, such as whether or not any articles are written, the number of articles written, or an article's "prominence," (defined variously as the total number of words, column length, or where the article appears within the paper). However, several important recent papers have instead applied content analysis to understand the qualitative nature of homicide coverage (22), focusing on whether offenders and victims are described in positive or negative terms (23–26) or coding articles according to a strict standardized criterion, such as the number of close relations of the victim (e.g., family, friends, etc.) mentioned in the article (26). In terms of the demographic groups considered, the most common finding in this literature has been reduced coverage of black victims. In terms of data availability, existing studies usually attempt

[1]To whom correspondence should be addressed. E-mail: ocasio.emi@gmail.com

to match homicide case reports to specific news articles by searching for victim or offender names based on a single large metropolitan area and its corresponding major newspaper; what geographic, temporal, and homicide-specific controls are able to be included in the analysis varies significantly across studies. We build on this literature along all three dimensions of analysis described above.

A central contribution of our analysis concerns our approach to measuring coverage, and specifically how we determine whether a news article humanizes a victim. Our approach takes a comprehensive view of the textual content of the articles in our sample rather than focusing on a single dimension of humanizing language by applying natural language processing (NLP) techniques to classify each article as humanizing or impersonal with respect to a particular victim. Specifically, our analysis uses GPT-3, a large language model (LLM) that has been pre-trained with 300 billion tokens on 45 terabytes of data obtained from a variety of sources including books and public web pages (27). NLP has shown utility in automating and standardizing content analysis (28–30). The value of language models in text classification lies in their ability to derive meaning out of precise textual context (31), thus affording a great deal of flexibility in determining where or not language is humanizing.[*]

A second central contribution of our analysis concerns how we understand the role of demographics in governing media coverage: we focus on the importance of understanding differential media coverage of crime through an intersectional lens. To do this, our regression analysis explicitly accounts for how race, age, and sex all *interact* to shape media coverage. Put differently, our analysis allows us to observe the effect of race on media coverage within age-sex cells, and thus paints a more nuanced picture of how media coverage differs across groups. This is important both because it avoids reducing the experience of different groups as discussed above, but also because it provides much richer empirical discipline needed to discriminate between various sociological theories of media coverage and crime.

Finally, our analysis is built around a new dataset that we construct from the FBI's Supplementary Homicide Reports (SHR) database and the universe of *Boston Globe* articles between 1976 and 1984. This allows us to include nearly all homicides in an entire state, and to match those homicides to nearly all corresponding articles in a major regional newspaper. Furthermore, in addition to the demographic information described above (both for victim and offender), the FBI's SHR data includes detailed information on the circumstances of each homicide, as well as when and where it occurred. This allows us to control for a wide range of factors that might otherwise confound our analysis of how demographics affect media coverage of crime.

Study design

Overview. We examine demographic correlates in newspaper coverage of homicides occurring in Massachusetts between 1976 and 1984. Homicide data, including date as well as demographics of both victims and suspected offenders, when known, is obtained from the publicly available FBI crime data. We match individual homicide reports to articles published in the *Boston Globe* morning editions about each homicide and determined whether each resulted in one or more published articles. Using NLP, we determine whether the newspaper articles included humanizing details about the victims or simply provided basic information in an impersonal way. Generalized linear models were applied to determine correlations between single and combined demographic factors (race, age, and sex of victims and/or offenders) and the amount and quality of coverage.

Homicide data and news articles. The universe of homicides available for our study includes all homicides committed anywhere in Massachusetts in the years 1976-1984. Rather than relying on police records from individual localities, we utilize data from the SHR (32), which captures homicide data from all local and state jurisdictions. The location and particular years were selected since they represent the confluence of dates between when case-level information is published in the SHR (starting in 1976) and when the digitized textual transcripts of the article content of a major regional newspaper (*The Boston Globe*) were available in bulk format for data mining (through 1984).

Table 1. Summary statistics by age, sex, and race

	Male			Female		
	White	Black	Other	White	Black	Other
Victim Age						
$\leq$18						
Hum. lang.	0.70	0.56	·	0.85	0.91	0.67
Artic./hom.	3.14	1.67	·	5.76	10.00	1.67
Homicides	84	33	0	68	15	3
18-29						
Hum. lang.	0.47	0.27	0.67	0.68	0.45	1.00
Artic./hom.	2.29	1.19	2.12	3.77	2.16	3.50
Homicides	339	168	8	114	44	2
30-49						
Hum. lang.	0.59	0.34	1.00	0.78	0.80	0.75
Artic./hom.	3.23	1.10	7.38	2.52	2.16	1.75
Homicides	337	126	8	79	25	4
50-69						
Hum. lang.	0.61	0.38	1.00	0.63	0.75	1.00
Artic./hom.	1.45	1.22	4.00	1.33	0.86	7.00
Homicides	143	36	2	64	7	1
$\geq$70						
Hum. lang.	0.60	1.00	·	0.65	1.00	·
Artic./hom.	1.06	3.80	·	1.09	2.00	·
Homicides	32	5	0	43	2	0

Hum. lang.: Share of homicides using humanizing language.
Artic./Hom.: Average articles per homicide.
Homicides: Number of homicides.

Our analysis includes each of the 1,800 Massachusetts homicides reported in the SHR as well as every newspaper article published in the *Boston Globe* for that time frame (1,215,056 articles, including non-news items such as advertisements and stock quotes). We develop and test a filtering algorithm to automate the identification of Massachusetts homicide (MH) articles, based on keyword/phrase matching of language related to homicides, Massachusetts locations and neighborhoods, and article types. The algorithm is calibrated to identify all possible MH candidates even though some of the articles initially

[*]A language model predicts the most likely next words to follow given a previous text, based on examples given during the model's training phase.

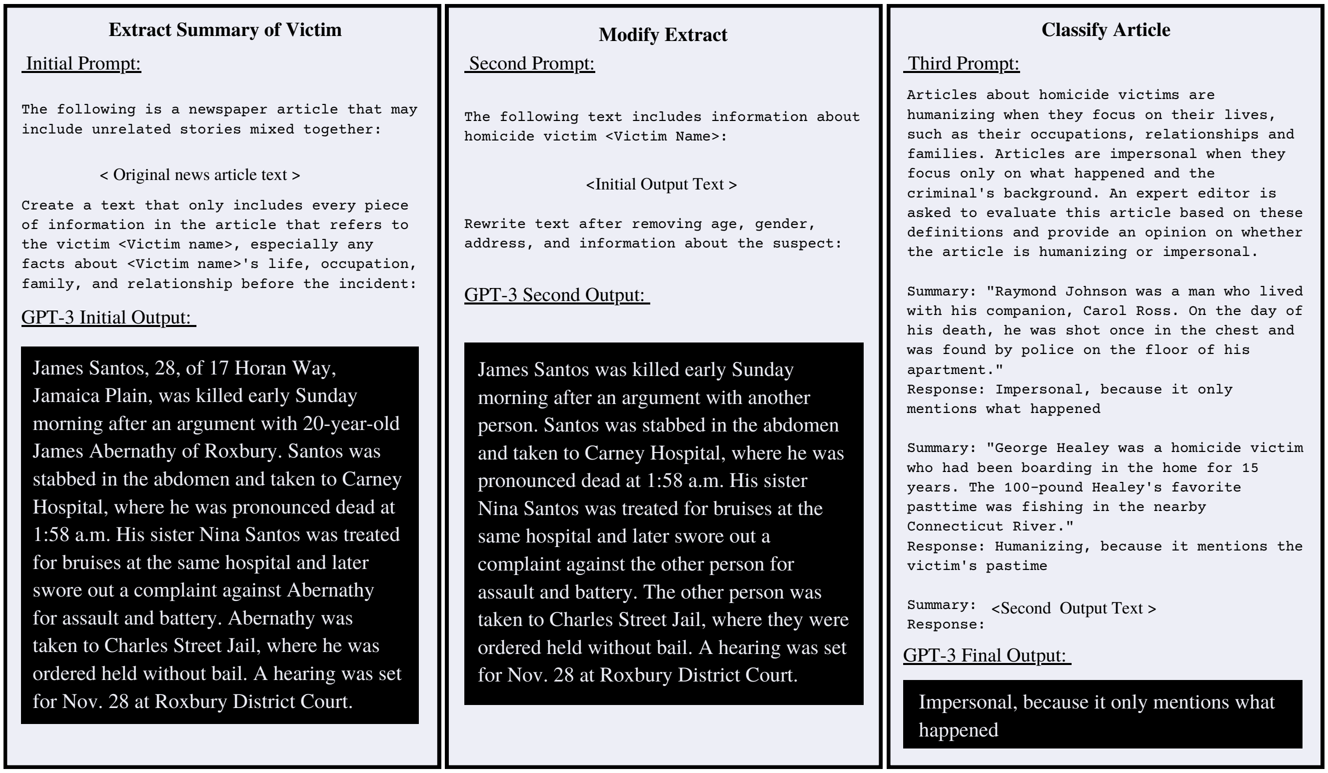

Fig. 1. GPT-3 prompt text and sample responses. Text in black background is actual response from GPT-3 for each of the three prompts. <Original news article text> is replaced with the full text of the news article. <Victim name> is replaced by the victim's full name. <Initial Output Text> in the second prompt is replaced by the full text response of the first prompt. <Second Output Text> in the third prompt is replaced by the full text response of the second prompt.

identified may be false positives due to the use of similar language, requiring further manual filtering to finalize the MH set. We then test the final algorithm against a randomly selected set of articles labeled as MH or not MH by a human reader, demonstrating 100% sensitivity and 98.2% specificity. The algorithm identified 22,640 of the 1.2 million articles as being potential MH, and subsequent manual inspection reduced the potential set down to a final confirmed set of 5,042 MH articles.

Using the demographic and crime detail information in the articles, we then manually match each one of the confirmed MH articles to one or more MHs from SHR. 835 articles that referred to pre-1976 MH are removed from further consideration. Of the remaining articles, 988 are unable to have any homicide record assigned to them because they represent homicides that were clearly not in SHR database, or the article is so vague in details that there is not enough information to assign to a particular homicide. The final count of articles positively matched and assigned to one or more SHR entries is 3,219, with a total of 1,071 homicides matched to one or more articles.

Humanizing coverage. We consider a homicide as being covered in a humanizing manner when at least one news article mentions additional information about the victim beyond age, race, location, and circumstances of death, resulting in the victim being presented as a person, not just a statistic. These details usually include information about the victim's occupation, family, background, interests, and/or personal comments about the victim from family, friends and neighbors. Impersonal (not humanizing) coverage of the homicide is when only the basic facts of the victim are presented and no additional background is included.

GPT-3 is used to automate the determination of whether an article is humanizing or impersonal in its coverage of the homicide victim. Our prompt strategy involves the use of three consecutive prompts to the GPT-3 text completion interface (Figure 1). An initial prompt sends the full article text to GPT-3 such that the completion text returned is an extract of the article containing exclusively the information about the particular victim in question. The second prompt submits this extract back with a request to summarize it by removing demographic information such as age and race of the victim, as well as information about the offender and the circumstances of the crime itself. The summary returned as completion text is more precisely focused on personalized victim information. The third prompt sends the summary from the second requesting a final classification as humanizing vs. impersonal using the few-shot technique. To determine whether an article is humanizing, we submit the three prompts to GPT-3 and use the first word (*humanizing* or *impersonal*) of the final response as the result. For each homicide victim, we consecutively send each article on that homicide to GPT-3 as outlined above, until at least one article is determined to be humanizing or all articles are considered impersonal.

We test the accuracy of GPT-3 humanizing determination against a human reader's ground truth labeling of 100 articles covering 30 randomly selected homicides, with GPT-3 agreeing with the human reader for 28 of the homicides (93.3% accuracy). Out of the 1,071 homicides with news articles evaluated for humanization, 602 homicides are classified as humanizing (56%).

Results

Table 1 reports summary statistics for the combined dataset. Specifically, within victim age-sex-race cells, the table reports (i) the share of victims described with humanizing language (Hum. lang.), (ii) the average number of articles per homicide (Artic./hom.), and (iii) the number of homicides (Homicides).

The goal of our empirical analysis is to use the data described above to understand how demographic characteristics influence the extent to which media coverage conveys homicide victims in a humanizing light. Specifically, the detailed homicide data from the SHR enables us to study rich *interactions* among demographic factors. It also allows us to control for demographic-correlated factors related to the nature of the homicide that would otherwise be likely to confound an analysis that only included demographic factors. The humanizing score that we construct with GPT-3 enables us to study how the demographic factors interact to shape humanizing coverage.

The analysis that follows first describes our empirical strategy and then presents three sets of results, each corresponding to a different measure of humanizing language: Our main results are based on a *composite humanizing score*, which measures the use of humanizing language in the full sample of homicides, including those receiving no coverage (which are deemed to be non-humanizing) and those receiving coverage (which we categorize as humanizing or non-humanizing as described above). Next, we consider an *extensive humanizing score*, which measures the quantity of coverage—i.e., the number of articles written about a homicide, irrespective of the nature of the language used in those articles. Finally, we consider a *conditional humanizing score*, which instead measures the quality of coverage—i.e., the likelihood of humanizing language, conditional on a homicide receiving coverage.

Empirical strategy. Formally, let h_v denote any of the three measures described above of the extent to which coverage of victim v is humanizing, and let X_v denote the corresponding vector of demographic and other controls from the FBI's SHR database, including:

1. Victim demographic controls: Race, sex, and age;
2. Offender demographic controls: Race, sex, age;[†]
3. Homicide controls: Weapon, relationship between victim/offender, specific circumstances of the homicide (e.g., gang-related), whether there were multiple victims, and whether there were multiple offenders;
4. Time (month-by-year) and county fixed effects.

For all three measures of humanizing coverage, we estimate the relationship between h_v and X_v within a generalized linear regression framework, i.e.

$$E[h_v|X_v] = f^{-1}(X_v'\beta)$$

where β is the vector of coefficients we seek to estimate and f is the link function mapping X_v to the conditional mean of h_v.

[†] We include an indicator for whether or not a murder was solved, which is implicitly an offender demographic control because, in that case, no offender demographic information is available.

Composite humanizing score. Our main results are based on the *composite humanizing score* ($h_v = h_v^{\text{comp}}$), which we define as an indicator variable that takes on a value of one if at least one article using humanizing language is written about homicide victim v and zero otherwise:

$$h_v^{\text{comp}} = \begin{cases} 0 & \text{No humanizing articles (incl. no articles)} \\ 1 & \text{At least one humanizing article.} \end{cases}$$

Thus, there are two ways in which a homicide victim can fail to be humanized: Either no article is written, or at least one article is written but none uses humanizing language.

Table 2. Determinants of Humanizing Coverage of Homicide

	Lin. Prob. Model		*Probit Model*	
	Dem.	Full	Dem.	Full
Victim				
Black	-0.18***	-0.15***	-0.53***	-0.53***
	(0.03)	(0.03)	(0.10)	(0.11)
Female	0.13***	0.17***	0.35***	0.61***
	(0.03)	(0.03)	(0.07)	(0.10)
Offender				
Black	0.08**	0.04	0.23**	0.16
	(0.04)	(0.04)	(0.11)	(0.13)
Female	-0.04	0.03	-0.13	0.10
	(0.04)	(0.04)	(0.13)	(0.16)
Controls/FE				
Unsolved	0.12	-0.01	0.44	-0.05
	(0.15)	(0.17)	(0.59)	(0.61)
Mult. victims		0.25***		0.86***
		(0.05)		(0.17)
Mult. offenders		0.04		0.17
		(0.04)		(0.14)
Weapon (FE)		×		×
Circum. (FE)		×		×
Relation (FE)		×		×
County (FE)		×		×
Time (FE)		×		×
Observations	1683	1683	1680	1629
Adjusted R^2	0.032	0.168		

Robust standard errors in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$
FE indicates inclusion of fixed effects for the corresponding variable.
Dem. = Model estimated only with demographic controls.
Full = Model estimated with additional non-demographic controls.
The sample consists of all homicides with a black or white victim under 70 years of age (other demographic cells did not contain enough observations for identification). Sample sizes differ across specifications because, as is well known, perfect predictors can lead to numerical problems in maximum likelihood estimation. As a result, these predictors and their associated observations are automatically dropped in the estimation procedure (as is the default behavior of most statistical packages).

We begin with a simple analysis in which we only consider a subset of demographic controls—victim/offender race and sex—thus neglecting age as well as potentially important interactions among these variables, to which we return below. Table 2 reports the results from this analysis. The first two columns correspond to estimation of a simple linear probability model ($f^{-1}(X_v'\beta) = X_v'\beta$) to facilitate interpretation

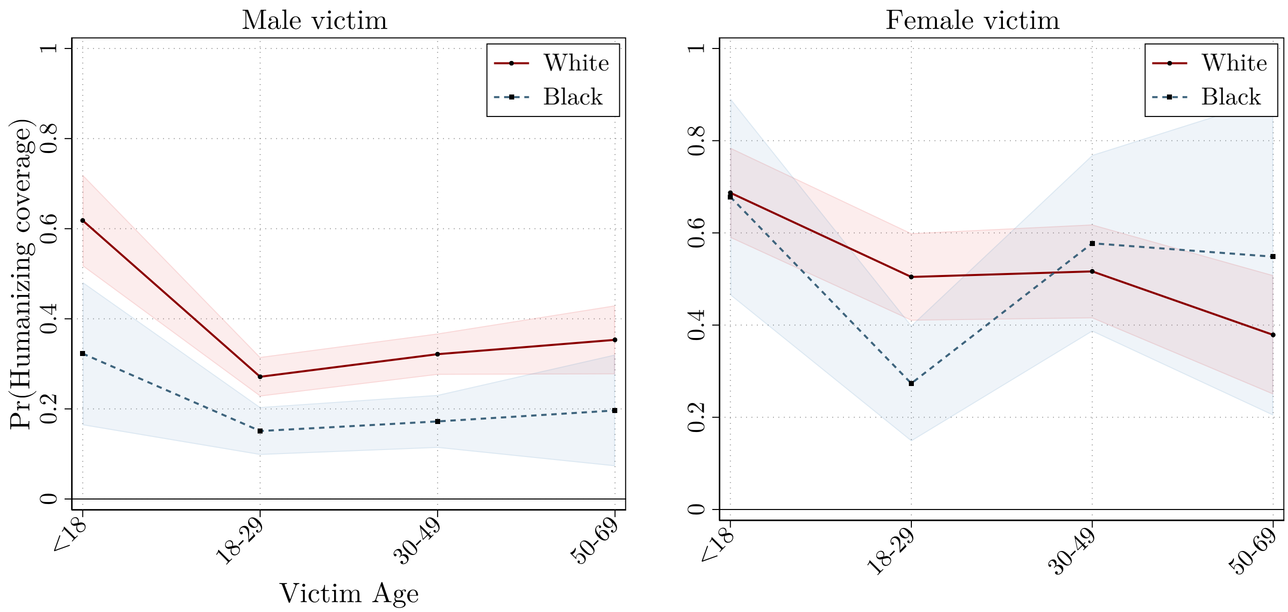

Fig. 2. Predicted probabilities of humanizing coverage by race, age, and sex with 95% confidence intervals. Results are computed from a Probit regression model including fully interacted demographic indicators as well as weapon, circumstance, relationship, county, and time fixed effects. See Table XYZ in the SI Appendix for the point estimates used to computed the predictive margins.

of parameters, and the second two columns correspond to estimation of a Probit model ($f^{-1}(X_v'\beta) = \Phi(X_v'\beta)$, where Φ is the cumulative distribution function of the standard normal distribution). Within each functional form specification, we report results when the model is estimated only with demographic controls as well as with the full set of controls listed above.

The results reveal three important insights into how humanizing coverage varies across demographic groups. First, referring to the results from the linear probability model, on average black victims are between 15 and 18 percentage points less likely to receive humanizing coverage than white victims (first row). Second, on average female victims are between 13 and 17 percentage points more likely to receive humanizing coverage than male victims (second row). Both results are significant at all conventional levels and robust to the inclusion of a rich set of controls and fixed effects. Third, in the models that fail to control for non-demographic factors (first and third columns), there is a significant positive relationship between an offender being black and the amount of coverage of homicide—homicides committed by black offenders appear to be roughly eight percentage points more likely to be described in a way that humanizes the victim than those committed by white offenders. However, when we account for the rich set of homicide-specific controls included in the FBI data, as well as time and locality fixed effects, this effect shrinks dramatically in magnitude and ceases to be distinguishable from zero. This observation highlights the importance of controlling for non-demographic factors that are likely to be correlated across demographic groups when analyzing the determinants of media coverage of crime.

While the analysis in Table 2 sheds important light on some demographic determinants of humanizing coverage of homicide, it fails to account for the rich ways in which intersectional demographic factors *interact* to shape that coverage. Accordingly, we re-estimate the fourth specification in Table 2, augmented with fully-interacted indicators for race, age, and sex.[‡] To facilitate exposition of the results, Figure 2 plots the predicted probabilities of humanizing coverage disaggregated by race, age, and sex. Specifically, the left panel plots the predicted probability of humanizing coverage for white and black males by age, and the right panel does the same for females.[§] Table 3 reports formal tests of equality of the predicted probabilities for black and white victims within each age-sex cell.

The results in Figure 2 reveal a much richer and more nuanced picture of how demographics interact to shape media coverage of crime compared with Table 2. First, among male victims, the probability of humanizing coverage is significantly higher for white victims than for black victims in all age

Table 3. Testing equality of probability of humanizing coverage between black and white victims

Age	Male		Female	
	Δ (Black - White)	p-value	Δ (Black - White)	p-value
<18	-0.30	0.00	-0.01	0.94
18-29	-0.12	0.00	-0.23	0.00
30-49	-0.15	0.00	0.06	0.58
50-69	-0.16	0.04	0.17	0.36

H_0: No black-white difference within age-sex cells ($\Delta = 0$).

[‡] Throughout the analysis, we focus on black and white victims under the age of 70 and partition victims and offenders into four age groups: children (< 18), young-adult (18-29), middle-aged (30-49), and older (50-69). Other groups do not contain enough observations for identification.

[§] Complete regression results are available upon request. Predictive margins are computed by integrating over all non-demographic covariates distributed as observed in the data.

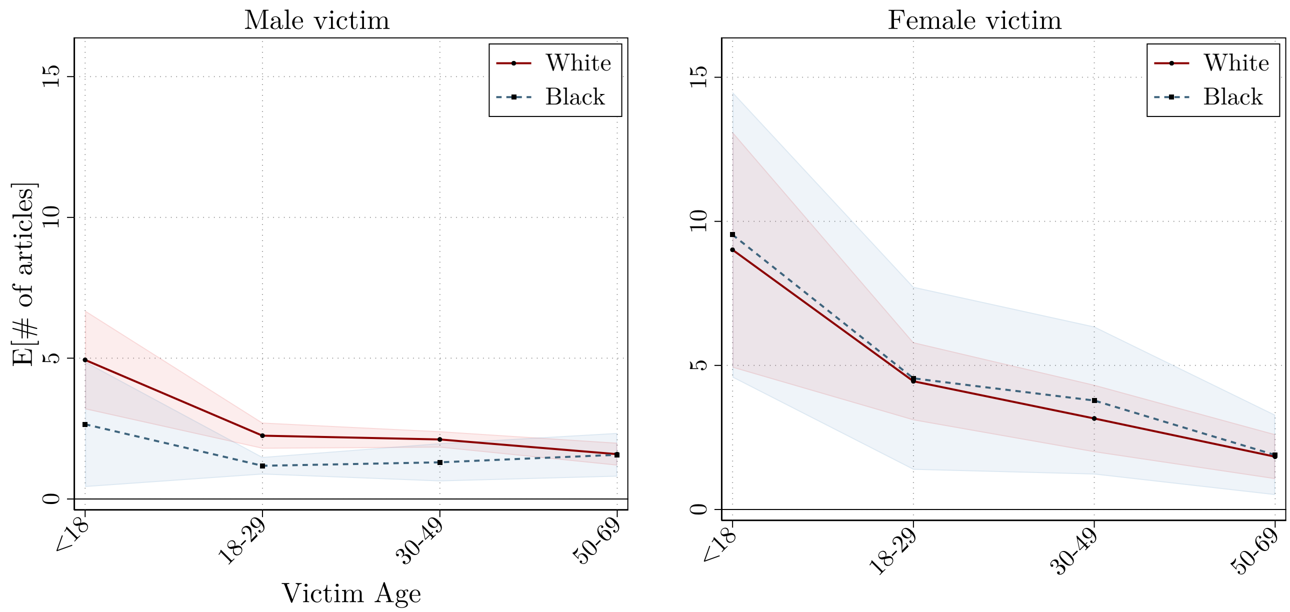

Fig. 3. Predicted number of humanizing articles by race, age, and sex with 95% confidence intervals. Results are computed from a Poisson regression model including fully interacted demographic indicators as well as weapon, circumstance, relationship, county, and time fixed effects. See Table XYZ in the SI Appendix for the point estimates used to computed the predictive margins.

categories, with the greatest difference among juveniles (30 percentage points) and an average difference of between 12 and 15 percentage points for all other age groups. As may be seen in Table 3, all of these differences within age groups are statistically significant at the 5% level, and all but the difference within the 50-69 year old category are statistically significant at the 1% level.[¶] Second, the same pattern does not hold for females: in general, there is not a statistically significant difference between white females and black females. However, a difference does emerge among young-adult females, in which case white victims are 23 percentage points more likely to receive humanizing coverage than black females. Third, females generally receive more coverage than their male counterparts, regardless of age. Finally, young males (18-29) of both races are less likely to be covered in a humanizing light than their older or younger counterparts.

The dependent variable in our analysis to this point—the composite humanizing score—implicitly reflects two ways in which media coverage can humanize a victim: through the quantity of coverage (e.g., whether or not an article is written about a homicide irrespective of the language of the reporting) and through the quality of that coverage conditional on an article being written. Below, we study these two sources of humanizing coverage in greater detail.

Extensive humanizing score. Because we have linked homicides from the FBI's SHR database to *all* corresponding *Boston Globe* articles covering those homicides between 1976 and 1984, we are able to measure the number of articles written about any given homicide. This represents a natural alternative measure of humanizing media coverage–over the course of multiple

Table 4. Testing equality of number of articles written about black versus white victims

Age	Male		Female	
	Δ (Black - White)	p-value	Δ (Black - White)	p-value
<18	-2.29	0.15	0.52	0.88
18-29	-1.07	0.00	0.10	0.95
30-49	-0.81	0.03	0.62	0.68
50-69	-0.02	0.96	0.06	0.94

H_0: No black-white difference within age-sex cells ($\Delta = 0$).

articles, a richer and more complex picture can be painted of a victim than would be possible in a single article (or, indeed, no articles). Accordingly, we define the *extensive humanizing score* ($h_v = h_v^{\text{ext}}$) as the number of articles written about a given homicide,

$$h_v^{\text{ext}} = \#\text{ of articles}$$

and estimate a Poisson regression ($f^{-1}(X_v'\beta) = \exp(X_v'\beta)$) where the dependent variable is now h_v^{ext}. The set of explanatory variables X_v is identical to those used in the analysis in Figure 2. Figure 3 summarizes the predicted number of articles written, broken down again by race, sex, and age, and Table 4 reports formal tests of equality of the predicted number of articles for black and white victims within age-sex cells.

With this alternative measure of humanizing coverage, we observe some similarities as well as some important differences relative to the results described above. Among the important similarities are systematically more coverage of white victims than black victims among males, virtually no racial discrepancy at any age among females, and the tendency for young black males (18-29) to receive less coverage than all other groups.

[¶]Note that overlapping confidence intervals in Figure 2 need not indicate a statistically insignificant difference. See the case of 50-69 year old males for a case in which the 95% confidence intervals overlap but in which the difference is significant at the 5% level.

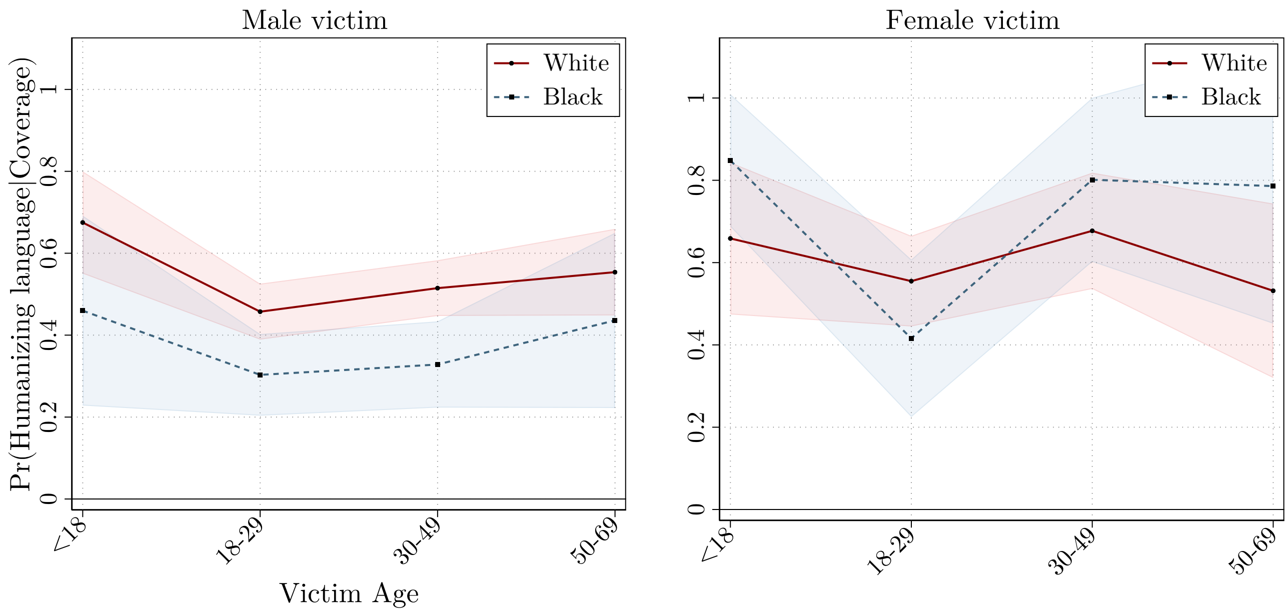

Fig. 4. Predicted probability of at least one humanizing article, conditional on at least one article being written, by race, age, and sex with 95% confidence intervals. Results are computed from a Probit regression model including fully interacted demographic indicators as well as weapon, circumstance, relationship, county, and time fixed effects. See Table XYZ in the SI Appendix for the point estimates used to computed the predictive margins.

However, when the dependent variable is the number of articles, the gap between males and females is substantially larger than previously. Moreover, there is no gap between white and black coverage of females at any age group, whereas the difference between white and black males is only robustly significant among young (18-29) and middle-aged (30-49) males (second and third rows of Table 4). The gap between black and white male children is once again the most pronounced, but the difference in this case is not significant (first row of Table 4).

Conditional humanizing score. Our third and final approach to understanding what determines humanizing language focuses strictly on the language used in the articles rather than the probability of coverage or the number of articles. Accordingly we define the *conditional humanizing score* ($h_v = h_v^{\text{cond}}$) as the probability of humanizing language conditional on at least one article being written, i.e.,

$$h_v^{\text{cond}} = \begin{cases} 0 & \text{No humanizing articles} \mid \geq 1 \text{ article} \\ 1 & \text{At least one humanizing article} \mid \geq 1 \text{ article} \end{cases}$$

and estimate a Probit model ($f^{-1}(X_v'\beta) = \Phi(X_v'\beta)$) where now the dependent variable is h_v^{cond} and we thus restrict attention to only those homicides about which at least one article was written. Furthermore, in addition to the standard set of controls used in Figures 2 and 3, we also control for the total number of articles written about a homicide, which will tend to be mechanically correlated with our dependent variable due to our approach to coding humanizing articles (see Materials and Methods section for details). Figure 4 reports the results from this analysis (analogously to Figures 2 and 3) and Table 5 reports formal tests of equality of the predicted probabilities for black and white victims within age-sex cells (analogously to Tables 3 and 4).

Table 5. Testing equality of probability of humanizing coverage (among homicides receiving coverage) between black and white victims

Age	Male		Female	
	Δ (Black - White)	p-value	Δ (Black - White)	p-value
<18	-0.22	0.11	0.19	0.12
18-29	-0.15	0.02	-0.14	0.22
30-49	-0.19	0.00	0.12	0.33
50-69	-0.12	0.33	0.25	0.21

H_0: No black-white difference within age-sex cells ($\Delta = 0$).

The results in Figure 4 and Table 5 add yet another layer of nuance to our understanding of the determinants of humanizing coverage. Specifically, we now observe patterns that are qualitatively quite similar to the patterns that we observed in Figure 2: A marked gap between humanizing coverage of black and white male victims with young males (18-29) receiving less coverage than other age groups and no universal difference between the language used to describe black and white female victims. Notwithstanding this second observation, and also similar to Figure 2, we see a pronounced racial gap among young (18-29) females, although in this case the gap is not statistically significant whereas it was in Figure 2. On the other hand, the gap between males and females of all ages is somewhat diminished relative to what we see in Figure 2, and the pervasive black-white gap among males that was significant for all age groups in Figure 2 is now only robustly significant among young (18-29) and middle-aged (30-49) males, similar to the results in Figure 3 and Table 4.

Discussion

Our study provides empirical evidence to support the hypothesis that newspaper depictions of homicides disproportionately humanize victims of certain demographic groups more than others. Our data does not support the idea that newsworthiness is mostly explained by an event's rarity. Although the observed greater degree of humanizing coverage of juvenile victims can be understood as a natural consequence of the relative infrequency of juvenile homicides, this explanation does not account for what we observe when comparing juvenile homicide coverage of young males by race: white male juveniles are much more likely to be humanized than black male juveniles despite the fact that there were almost three times as many white juvenile victims as there were black juvenile victims.

Differences across demographic groups are brought into sharp contrast when measured intersectionally. Across age groups, we observe significantly smaller humanizing rates for black male victims as compared to white males, while female white victims overall are not more likely to be humanized than female black victims. However, when we account for differential effects across age groups, significant differences are seen between certain age groups: black women between 18 and 29 are significantly less likely to be humanized than white women of the same age. Moreover, while black men are humanized less across all age groups, the gap is largest for juveniles.

These findings are consistent with ideal victim theory: young people are seen as more vulnerable, while being a black victim counteracts this by being seen as aggressive or strong. While age, sex, and race all influence societal perceptions of innocence and vulnerability, they act in different directions across groups. Our results can be understood as reflecting identified stereotypes towards black women and men; younger black women, for example, are masculinized and seen as strong, falling into the "Jezebel" or "strong black woman" trope, while older black women instead fall into the "mammy" stereotype which is associated with caring and motherhood. Similarly, for males, the largest difference in the humanizing probability is between black and white juveniles. This is the most vulnerable and innocent age group, so the observed difference is consistent with dehumanizing perceptions of black boys in particular as older, stronger, and less innocent (33).

Ideal victimhood thus provides the most consistent explanation of the humanization patterns we observe in our data. Children are the most ideal victims and thus they are the most likely age group to be humanized, followed by the elderly. Women are perceived as weaker, more innocent, and subsequently more ideal victims than men and thus they are more likely to be humanized. When additionally considering race, we see specific interactions that are also explained by ideal victim theory–younger adult black women and black males of all ages, but most particularly black juveniles, are less likely to be humanized because they are perceived as stronger, less vulnerable, and less innocent.

The results also demonstrate the importance of a methodological approach to media distortion analysis that combines the presence/extent of coverage (quantitative measures) with determination of humanizing language (qualitative measures). When we use the composite humanizing score, we detect significant differences between the humanizing probabilities of black and white women between 18 and 29. The extensive humanizing score regression, which instead considers article count, fails to detect this difference. The conditional humanizing score regression, which considers the humanizing categorization only among the subset of homicides about which at least one article is written, predicts a small difference which is not statistically significant.

Limitations. One potential limitation of our study is inherent in the lack of a source of crime reports that reliably includes 100% of the homicides in the analyzed location and time period. We thoroughly search all newspaper articles for MH-related stories independent of the crime reports, and subsequently find a number of cases not included in the SHR. This gap has not usually been identified or measured in media distortion analysis since the articles are usually found from a search by victim and/or offender name based on the available crime reports. Nevertheless, studies of SHR accuracy have validated their reliability when compared with aggregate data from the National Vital Statistics System, with the exception of cities with very low numbers of black homicides (34, 35). Specifically, some infrequent kinds of homicides, in particular justifiable homicides (e.g., law enforcement killings of felons during commission of a crime) show very low rates of inclusion in SHR (36).

The use of automated content analysis via GPT-3 allows for more scalable and consistently reproducible determination of humanizing article language, especially when analyzing large numbers of articles. Although it eliminates human variance and coding error, it may introduce biases arising from the language model itself. This potential for bias has been recognized by the developers of GPT-3 (27). When GPT-3 was provided racial cues during prompting, the developers noticed the resulting text differed in sentiment based on the prompted race. This was most likely due to the inherent discrepancy in depictions of race among the millions of source articles used to train the model. Our use of GPT-3 does not explicitly provide race in the prompts and also removes whatever references to demographic characteristics may be in the article during the second prompt (Figure 1), so the final determination of humanization is as race- and age-agnostic as possible. Gender is harder to remove completely from the prompts because of the associated pronouns and the presence of gender-specific names (this may also be an issue with race to a much smaller extent). In addition to making the prompts neutral, we also do not ask GPT-3 to generate new language, only to summarize and extract information from existing articles, as well as to make a final determination. Future studies should explore the effects of artificially adding, removing, or changing racial depictions in prompts on GPT-3 humanization classification accuracy.

Finally, we emphasize that our estimates do not necessarily imply a causal relationship between a victim's demographics and humanizing coverage. While inclusion of the rich set of interacted demographic controls and non-demographic fixed effects made available in the SHR helps to account for the possibility of correlated confounding factors explaining the demographic patterns we observe in a way that goes beyond much of the existing literature, we cannot rule out that other such factors may be unobserved and driving the relationships we find in the data.

Conclusion

We document significant differences in both the quantity and quality of media coverage of homicide across victims from different demographic groups, with stereotypical ideal victims more likely to be humanized than those considered non-ideal victims. Average differences across races are large and pronounced, with black victims receiving significantly less coverage than white victims, consistent with historical stereotypes surrounding innocence and blackness. Yet once we allow for differential effects of race across age groups and sex, we find that the average effects of race conceal substantial heterogeneity across different sub-populations: For example, the fact that black victims receive less humanizing coverage than white victims on average is driven almost entirely by males; the only difference across races among females appears among 18-29 year-olds; and the racial gap among males is twice as large among juveniles as it is among older age groups.

Our findings thus demonstrate that failing to account for intersectionality in analyses of media coverage of crime both risks the erasure of important variation in human experience across groups and makes it difficult to discriminate between sociological theories. In addition, our findings demonstrate how content analysis (facilitated by the use of a state-of-the-art large language model, GPT-3) can be applied to produce holistic measures of the extent to which media language is humanizing and how that language is applied unequally. We hope that future researchers will learn from these insights and continue to work towards deepening our understanding of how the media both reflects and influences how society perceives members of different demographic groups.

Materials and Methods

Homicides. SHR data was downloaded from the Murder Accountability Project website (37). It includes individual entries for each homicide, with year and month of the homicide, victim demographics, offender demographics if there is a known offender, the police department that reported the case, and other information about the homicide such as the method of death (shooting, stabbing, strangulation, etc.), the offender relationship with the victim, the kind of weapon, and circumstances of the case (whether it was related to argument, robbery, drug-related, etc.). SHR does not include personal information on the victims or offenders, such as names or addresses.

News articles. We obtained the text of newspaper articles from the ProQuest Historical Newspapers database, which includes the digitized text of every *Boston Globe* article published through 1984, available in bulk for analysis via Text Data Mining Studio. For our analysis, we extracted the text of every article published in the morning edition for the years 1976 through 1984. These text transcripts were originally produced via optical character recognition from scanned paper copies of the newspapers, and as such they contained some errors such as misspelled words, occasional extraneous spaces within the words, uneven sentence formatting, and lack of paragraph separation. Despite these issues the article texts were readable and understandable. The article database also includes additional metadata about each article, including the date, title, and article type (e.g., news article, advertisement, editorial, stock quote).

Article filtering. Our automated filtering algorithm included three main criteria for determining whether an article may be MH: (1) the article type, (2) presence of keywords or word combinations related to homicides, (3) reference to Massachusetts locations. The article type criteria removed from consideration database entries such as classified ad copy and paid birth/death notices. The algorithm looked for the presence of homicide-related keywords (e.g.: "murder," "slain") as well as word combinations (e.g.: "shot to death"). It also ensured the articles referred to homicides in Massachusetts by ensuring reference to a location from a list of towns, cities, and counties in Massachusetts as well as neighborhood and places in Boston.

Our goal for the algorithm was to ensure it did not miss any MH articles (false negatives) while reasonably minimizing the number of non-MH articles it erroneously marked as MH (false positives). We randomly selected article subsets to measure the algorithm's accuracy during the initial development round, three rounds of algorithm optimization validation, and a final testing round. Articles for each subset were assigned by randomly picking dates within the 1976-1984 time span and including in the subset all the articles for the selected days. Every article in all the subsets was ground truth labeled MH/not MH by a human reader, and sensitivity and specificity of the algorithm results against the ground truth were computed at the end of each round. The final algorithm specifications are described in SI Appendix.

Once the algorithm was finalized it was used to perform the first pass filter of the full article database. Articles identified as possible MH by the algorithm were then manually inspected by a human reader who further filtered the group down to the final number of validated MH articles used in the analysis.

Homicide matching to articles. Since names of victims and offenders are not included in the SHR, we used information in the news articles including date and location of homicide, victim and offender demographics, and other facts such as type of weapon, circumstances and victim-offender relationship, to identify each SHR homicide entry by victim name. Based on this we were able to determine whether each MH article matched to zero, one, or more SHR entries. The final multiplicity relationship between articles and homicides was zero to many in both directions.

Humanizing determination via NLP. Like other similar LLMs, GPT-3 is a pre-trained AI model and thus it does not require large training sets for classification. We used the few-shot prompt technique [ref] in which, using natural English text, a couple of examples are provided as a way to "prompt" the model to follow the pattern. To optimize the accuracy of our prompt design we manually adjudicated a random subset of homicides with their respective articles as humanizing or impersonal ground truths and used this validation set to tweak the prompts and improve the accuracy. Once all three prompts were deemed to be optimized we measured its performance against a randomly selected test set of homicides for which we manually determined ground truths.

The prompts with the accompanying article or extract text were submitted via the GPT-3 text completion interface (https://api.openai.com/v1/completions) using the text-davinci-002 model with the following interface parameters: temperature = 0, max_tokens = 256, frequency_penalty = 0, presence_penalty = 0. The temperature parameter in particular removes the element of randomness guaranteeing the same response after repeated identical prompts. The Davinci GPT-3 model has an inherent limit of the length of the text it can analyze. This limit of 4,000 tokens (each token is a meaningful part of a word, usually 750 words would average out to 1,000 tokens) includes both the prompt and the response. There were three homicides in the total set for which one of their respective articles was so long that we could not evaluate via GPT-3. For these three homicides we determined their humanizing classification by human inspection.

Regression analysis. The empirical framework for our regression analysis is described in the Results section above. Probit regressions are estimated in Stata 17 using the *probit* command and Poisson regressions are estimated in Stata 17 using the *ppmlhdfe* package (38). Complete replication files are available upon request.

1. SM Chermak, *Victims in the news: Crime and the American news media.* (Westview Press Boulder, CO), (1995).
2. DJ Paulsen, Murder in black and white: The newspaper coverage of homicide in houston. *Homicide Stud.* **7**, 289–317 (2003).
3. M Meyers, *News coverage of violence against women: Engendering blame.* (Sage Publications), (1996).

4. SE Clayman, A Reisner, Gatekeeping in action: Editorial conferences and assessments of newsworthiness. *Am. Sociol. Rev.* pp. 178–199 (1998).
5. DT Lowry, TCJ Nio, DW Leitner, Setting the public fear agenda: A longitudinal analysis of network tv crime reporting, public perceptions of crime, and fbi crime statistics. *J. communication* **53**, 61–73 (2003).
6. JG Boulahanis, MJ Heltsley, Perceived fears: The reporting patterns of juvenile homicide in chicago newspapers. *Crim. Justice Policy Rev.* **15**, 132–160 (2004).
7. MD Tans, SH Chaffee, Pretrial publicity and juror prejudice. *Journalism Q.* **43**, 647–654 (1966).
8. S Graham, BS Lowery, Priming unconscious racial stereotypes about adolescent offenders. *Law human behavior* **28**, 483–504 (2004).
9. KN Dukes, SE Gaither, Black racial stereotypes and victim blaming: Implications for media coverage and criminal proceedings in cases of police violence against racial and ethnic minorities. *J. Soc. Issues* **73**, 789–807 (2017).
10. N Christie, The ideal victim in *From Crime Policy to Victim Policy: Reorienting the Justice System*, ed. EA Fattah. (McMillan, London), pp. 17–30 (1986).
11. JA Lewis, JC Hamilton, JD Elmore, Describing the ideal victim: A linguistic analysis of victim descriptions. *Curr. Psychol.* **40**, 4324–4332 (2021).
12. M Wilson, E O'Brien, Constructing the ideal victim in the united states of america's annual trafficking in persons reports. *Crime, Law Soc. Chang.* **65**, 29–45 (2016).
13. J Bouchard, JS Wong, K Gushue, The differential portrayal of 'sympathetic'homicide victims in the media. *Int. Rev. Vict.* **26**, 314–331 (2020).
14. LJ Long, The ideal victim: A critical race theory (crt) approach. *Int. Rev. Vict.* **27**, 344–362 (2021).
15. D Steffensmeier, J Ulmer, J Kramer, The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology* **36**, 763–798 (1998).
16. JW Johnstone, DF Hawkins, A Michener, Homicide reporting in chicago dailies. *Journalism Q.* **71**, 860–872 (1994).
17. A Weiss, SM Chermak, The news value of african-american victims: An examination of the media's presentation of homicide. *J. Crime Justice* **21**, 71–88 (1998).
18. SB Sorenson, JG Manz, RA Berk, News media coverage and the epidemiology of homicide. *Am. journal public health* **88**, 1510–1514 (1998).
19. RJ Lundman, The newsworthiness and selection bias in news about murder: Comparative and relative effects of novelty and race and gender typifications on newspaper coverage of homicide in *Sociological forum*. (Springer), Vol. 18, pp. 357–386 (2003).
20. J Gruenewald, J Pizarro, SM Chermak, Race, gender, and the newsworthiness of homicide incidents. *J. criminal justice* **37**, 262–272 (2009).
21. J Schildkraut, AM Donley, Murder in black: A media distortion analysis of homicides in baltimore in 2010. *Homicide Stud.* **16**, 175–196 (2012).
22. D Humphries, Serious crime, news coverage, and ideology: A content analysis of crime coverage in a metropolitan paper. *Crime & Delinquency* **27**, 191–205 (1981).
23. MH Barlow, DE Barlow, TG Chiricos, Economic conditions and ideologies of crime in the media: A content analysis of crime news. *Crime & Delinquency* **41**, 3–19 (1995).
24. D Pritchard, Race, homicide and newspapers. *Journalism Q.* **62**, 500–507 (1985).
25. RM Entman, Blacks in the news: Television, modern racism and cultural change. *Journalism quarterly* **69**, 341–361 (1992).
26. K White, F Stuart, SL Morrissey, Whose lives matter? race, space, and the devaluation of homicide victims in minority communities. *Sociol. Race Ethn.* **7**, 333–349 (2021).
27. T Brown, et al., Language models are few-shot learners. *Adv. neural information processing systems* **33**, 1877–1901 (2020).
28. Y Wang, J Tian, DS Ones, RN Landers, , et al., Using natural language processing and machine learning to replace human content coders. *Psychol. Methods* (2022).
29. CH Brown, , et al., Ph.D. thesis (2018).
30. R Rezapour, Ph.D. thesis (University of Illinois at Urbana-Champaign) (2021).
31. JL Jensen, et al., Language models in sociological research: An application to classifying large administrative data and measuring religiosity. *Sociol. Methodol.* **52**, 30–52 (2022).
32. WC Regoeczi, JP Jarvis, Fbi supplementary homicide reports. *The Encycl. Res. Methods Criminol. Crim. Justice* **1**, 43–46 (2021).
33. PA Goff, MC Jackson, BAL Di Leone, CM Culotta, NA DiTomasso, The essence of innocence: consequences of dehumanizing black children. *J. personality social psychology* **106**, 526 (2014).
34. C Loftin, D McDowall, K Curtis, MD Fetzer, The accuracy of supplementary homicide report rates for large us cities. *Homicide studies* **19**, 6–27 (2015).
35. WM Rokaw, JA Mercy, JC Smith, Comparing death certificate data with fbi crime reporting statistics on us homicides. *Public Heal. Reports* **105**, 447 (1990).
36. BK Finch, et al., Assessing data completeness, quality, and representativeness of justifiable homicides in the fbi's supplementary homicide reports: A research note. *J. Quant. Criminol.* **38**, 267–293 (2022).
37. T Hargrove, Murder accountability project (2022).
38. S Correia, P Guimaraes, T Zylkin, Ppmlhdfe: Stata module for poisson pseudo-likelihood regression with multiple levels of fixed effects. (2019).