

The New York Times distorts the Palestinian struggle: A case study of anti-Palestinian bias in US news coverage of the First and Second Palestinian Intifadas

Media, War & Conflict

2024, Vol. 17(1) 116–135

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17506352231178148

journals.sagepub.com/home/mwc**Holly M Jackson** 

Massachusetts Institute of Technology, USA and Institute for the Study of Human Rights, Columbia University, USA

Abstract

This article provides a methodologically novel, large-scale proof of historical bias against Palestine in a newspaper of international importance – *The New York Times* (NYT) – during the First and Second Palestinian Intifadas. Using state-of-the-art natural language processing toolkits as well as a regression model with over 90 percent accuracy based on a carefully validated word bank, the author analyzes over 33,000 NYT articles for (1) their use of active/passive voice, and (2) the objectivity, tone, and violent sentiment of the language used. She follows up her quantitative analysis with a qualitative validation step, analyzing biased articles in each period. In conjunction with historical context, the article shows that anti-Palestinian bias persisted disproportionately in the NYT during both periods and, in fact, worsened from the First Intifada to the Second. This work builds on a history of qualitative research on anti-Palestinian bias in the US media and attempts to provide a methodological contribution that encourages conversation between quantitative and qualitative metrics of bias.

Keywords

content analysis, Israel, Israeli–Palestinian conflict, machine learning, media bias, natural language processing, *New York Times*, Palestine

Introduction

In this study, I identify bias against Palestine in a newspaper of international importance – *The New York Times* (NYT) – as a case study in the scope of a larger problem of

Corresponding author:

Holly M Jackson, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA.

Email: holly_jackson@berkeley.edu

anti-Palestinian bias in US news coverage. I center my analysis on two important periods in the modern history of the Palestinian struggle – the First and Second Intifadas – both marked by periods of rapid change and an increase in global conversation about Palestine and Israel. From this analysis, I intend to, first, develop a well-informed metric for bias based on past qualitative work, second, provide a more rigorous computational analysis of rhetorical bias in the NYT than other content analysis studies have to date and, third, develop an accurate measure of how bias in the NYT evolved between the Intifadas.

Background

A primer on media bias

The NYT's ethical journalism handbook states, 'The goal of *The New York Times* is to cover the news as impartially as possible – "without fear or favor"' (*The New York Times*, nd). Fairness, impartiality, objectivity, neutrality, etc. are often claimed by large US newspapers – in opposition to the scandalous accusation of bias. Academic literature has analyzed the uncharacteristic devotion to objectivity in US journalism since its rise to prevalence in the late 18th/early 19th centuries. There is much debate as to the legitimacy of objectivity and the integrity of this movement.

Scholars like McQuail and Boudana strike a similar tone to the NYT's purpose statement. They speak highly of objectivity as a journalistic ideal and both believe it can be realistically approximated by an ever-growing list of strategies (Boudana, 2011; McQuail, 1992, 1994).

Many other scholars (arguably, a strong majority) see objectivity as a manufactured notion (Altheide, 1984; Altschull, 1995; Glasser, 1988; Hackett, 1984; Hémanus, 1976; Schudson, 2001; Tuchman, 1972). Schudson (2001) provides an historical analysis of the (very recent) origins of journalistic objectivity, framing the 'objectivity norm' as the inexact product of relativist cultural factors. Glasser (1988) and Hémanus (1976) both argue that the notion of journalistic objectivity constitutes bias in and of itself – as it treats facts only as true or false statements and neglects the impact of their framing. Hackett (1984: 236) agrees framing imposes partisanship, referring to language 'as a structuring agent, rather than a neutral transmission belt'.

Tuchman (1972), Altschull (1995), Altheide (1984), and Glasser (1988) see objectivity as a malleable tool – rather than a rigid philosophical notion – that was developed to safeguard journalists from criticism. Tuchman (1972: 676) coined the phrase 'strategic ritual' to describe the established procedures 'through which newsmen protect themselves from critics and lay professional claim to objectivity'. Altschull (1995), in particular, brought up the economic advantages newspapers reaped from proclaiming an objective approach.

While most scholars agreed objectivity is either practically unattainable or wholly fabricated, none denied the importance of characterizing journalistic bias. In constructing their varied analyses, many delineated useful metrics for bias – including diction, framing, selective reporting, and chronology (Hackett, 1984; McQuail, 1992). In addition, even in the few studies that considered objectivity an aspirational ideal, none argued objectivity has been achieved.

Identifying media bias remains an important task to challenge the purported objectivity of mass media in the USA. This study aims to build on a body of literature that contextualizes the dangers of denying the strong positionality of leading newspapers such as the NYT.

Anti-Palestinian bias in US media

Studies have shown there is clear Orientalist bias in mainstream US reporting on Palestine. In *The Israel Lobby*, Mearsheimer and Walt (2007: 169) argue, 'The American media's coverage of Israel tends to be strongly biased in Israel's favor.' Khalidi (2021) provides commentary on anti-Palestinian framing in US reporting (especially the NYT) in 'The Hundred Years' War on Palestine'. In addition, Bazian (2015) has analyzed anti-Palestinian bias as intrinsically connected to the Islamophobic campaigns in mainstream US media.

Many scholars have identified specific patterns of anti-Palestinian bias. In *Blaming the Victims* Said (2001) and Chomsky (2001) argue that news coverage of Palestinian violence is marked by Orientalist stereotype rather than evidence of humanitarian crimes committed. Chomsky (2001: 109) comments, 'As in the rule of properly sanitized history, Palestinians carry out terrorism, Israelis then retaliate, perhaps too harshly. In the real world, the truth is often rather different.' Chomsky explains that Israeli terrorism has rarely been criticized in mainstream media and is even sometimes celebrated. Meanwhile, Palestinians are blamed for instigating attacks without evidence or context provided (pp. 134–136). Similarly, Zelizer et al. (2002: 295) found that 'the Times tended to portray the Israelis as victims and the Palestinians as aggressors in its headlines', and Ackerman (2001) traced a 'systematic absence of context in U.S. reporting' on recent events in Palestine that obfuscates Israeli violence.

Nonetheless, much analysis of anti-Palestinian bias in US media is done outside the academy. Commentary sympathetic to the Palestinian struggle frequently results in intense backlash (Roy, 2010: 23) and, as a result, it is sometimes likely to see analyses performed by advocacy groups and non-profit organizations as well as individual whistleblowers on social media.

Existing literature on content analysis

While many scholars have called attention to patterns of bias, fewer researchers have done large-scale content analyses on bias against Palestine in the US media (for a European case study, see Sirhan's, 2021, book on anti-Palestinian bias in British news).

One example is a report published by 416Labs (Siddiqui and Zaheer, 2018) that used natural language processing (a machine learning technique) to analyze the diction in headlines of five major US newspapers (including the NYT) over a 50-year period. The authors tracked whether headline topics and news sources were more Palestinian or Israeli centric and found a strong bias towards Israelis (pp. 2–3). Another analytical work by *If Americans Knew* compares mentions of Palestinian and Israeli deaths in NYT articles during two periods in the Second Intifada. Overall, the report finds that Israeli deaths

are consistently over-reported while Palestinian deaths are consistently under-reported during the study period (Weir, 2005).

Frequently citing the *If Americans Knew* report, Caballero (2010) similarly quantified discrepancies in death statistics to establish bias on NYT coverage of Israel's 2008–2009 Operation Cast Lead. In addition, Al-Sarraj and Lubbad (2018) used supervised machine learning to identify pro-Israeli bias in reporting on the Gaza War in summer 2014, and Aziz (2007) found NYT and AP coverage of the 2006 Israeli invasion of Lebanon justified Israeli violence while condemning and exaggerating Palestinian violence. These findings serve as quantitative evidence of patterns of bias that have been established in previous qualitative studies of the same periods.

The organization *Fairness and Accuracy in Reporting* (FAIR, 2002) has also completed a number of reports on anti-Palestinian bias in US news, including an investigation of the disproportionate use of the word 'retaliation' to describe Israeli vs Palestinian violence (79% vs 9% in Second Intifada headlines). FAIR conducted a number of other short reports on pro-Israeli bias, including a 2009 report revealing only 2 percent of US news organizations had referred to the Israeli occupation using the phrase 'occupied territories' and a 2002 report exposing NPR for disproportionately reporting Israeli casualties (Abunimah, 2002; Ackerman, 2009).

Conflicting perspectives on the evolution of bias

It is clear from the secondary literature that there is bias against Palestine in US news coverage. However, it remains unclear how this bias is evolving. A few scholars have commented on changes in public opinion, but none have directly traced a change in news coverage. Sara Roy (2010: 25–27) hypothesizes an increase in public sympathy for Palestine and criticism of Israel since the early 2000s. She attributes the evolution, in part, to prominent Israeli and US figures denouncing Israeli humanitarian crimes and films that show a more realistic portrayal of the Israeli occupation.

In contrast, Rashid Khalidi (2021) argues that public opinion towards Palestine degraded from the First to Second Intifada. While Khalidi criticizes bias in news coverage of the First Intifada, he comments that 'its unifying effect and largely successful avoidance of firearms and explosives . . . helped to make its appeal widely heard internationally, leading to a profound and lasting positive impact on both Israeli and world public opinion' (p. 74). During the Second Intifada, Khalidi argues that public opinion of Palestine took an overwhelmingly negative turn due to an increase in coverage of Palestinian violence that eclipsed 'the much greater violence perpetrated against the Palestinians' (pp. 214–215). These sources leave us with a disjointed and incomplete understanding of the evolution of bias.

Methodology

While there is a body of existing literature that identifies bias against Palestine in US media and performs large-scale content analyses, few quantitative studies adapt metrics from past qualitative research or analyze article texts beyond headlines. While the literature shows there are discrepancies in the number of reports about deaths and attacks

(Weir, 2005), for the most part it does not address discrepancies in how this reporting is done. In addition, there is no clear consensus on how bias is evolving over time (Khalidi, 2021; Roy, 2010). These gaps in existing literature necessitate a thorough content analysis that reaches farther than metadata but also spans a period during which a change in bias (or lack thereof) is distinctly identifiable. In addition, there is a need for a large-scale quantitative study that thoughtfully builds on past qualitative work.

I use content analysis to focus on bias in the language of reporting over two distinct periods. Content analysis allows researchers to analyze larger quantities of material more efficiently (Grimmer and Stewart, 2013: 2). Combined with a supporting qualitative analysis, it allows breadth in my analysis of bias in the NYT. My analysis identifies two key features of NYT articles about Palestine and Israel: voice and tone.

First, I identify whether actions by Palestinian and Israeli groups are being described in the active or passive voice. For every verb, I identified the perpetrator and recipient of the action (i.e. whether they were a Palestinian or Israeli group or individual). An analysis of voice is common in other studies of bias because the passive voice, which is often discouraged in formal writing, allows the writer to de-emphasize or entirely omit the perpetrator of an action in a sentence (Frazer and Miller, 2008; Kazenin, 2008; Söğüt, 2018). We can consider excerpts from May 2021 NYT articles to see how voice plays a critical role in introducing bias (Cohen, 2021; Kingsley, 2021). In one article, the journalist uses the passive voice to describe the murder of Palestinians. '[M]ore than 230 Palestinians were killed in Gaza, including 67 children', he explains (Cohen, 2021). Because the author uses the passive voice, he never identifies the perpetrators. A reader knows Palestinians have died, but is left clueless as to who killed them. In contrast, another author writes, 'The rockets fired by Hamas and its Islamist ally, Islamic Jihad, killed at least six Israeli civilians, including a 5-year-old boy and one soldier' (Kingsley, 2021). The author uses the active voice in this example so that the reader can immediately identify the perpetrator as a Palestinian group. To identify similar instances of bias, I classified the voice of all sentences involving Palestinian or Israeli subjects in my dataset.

Second, I classify the objectivity and tone of language – especially that used to report on violence – based on pre-trained natural language processing (NLP) toolkits as well as my own regression model based on a carefully-validated word bank. Even when verbs are conjugated in the same voice, objectivity and tone can introduce bias. We can see an example of this within the same 2021 NYT article (Kingsley, 2021). The author writes, 'They [Jews] threw rocks at hotels housing Arabs, who hurled objects from their windows in return.' In this example, all verbs appear in the active voice. But, the author uses a more neutral verb ('throw') to describe Israeli violence and a less neutral verb ('hurl') to describe Palestinian violence. I incorporated a classification of objectivity and tone into my analysis to reveal similar examples.

All my content analysis was performed in Python using state-of-the-art natural language processing libraries (Keras, nd; NLTK 3.6.2, nd; scikit-learn, nd; spaCy, nd). Natural language processing can identify syntactical and semantic aspects of language on a large scale. In order to analyze sentence voice, I utilized part-of-speech tagging and parse tree toolkits from spaCy and NLTK to identify sentence subjects and their associations as well as distinguishing between active and passive voice (Jagota, 2020). To analyze tone, I used built-in tools to classify associated sentiment of words throughout the corpus of articles on Palestine and Israel. These scores are assigned using the NLTK interface with SentiWordNet

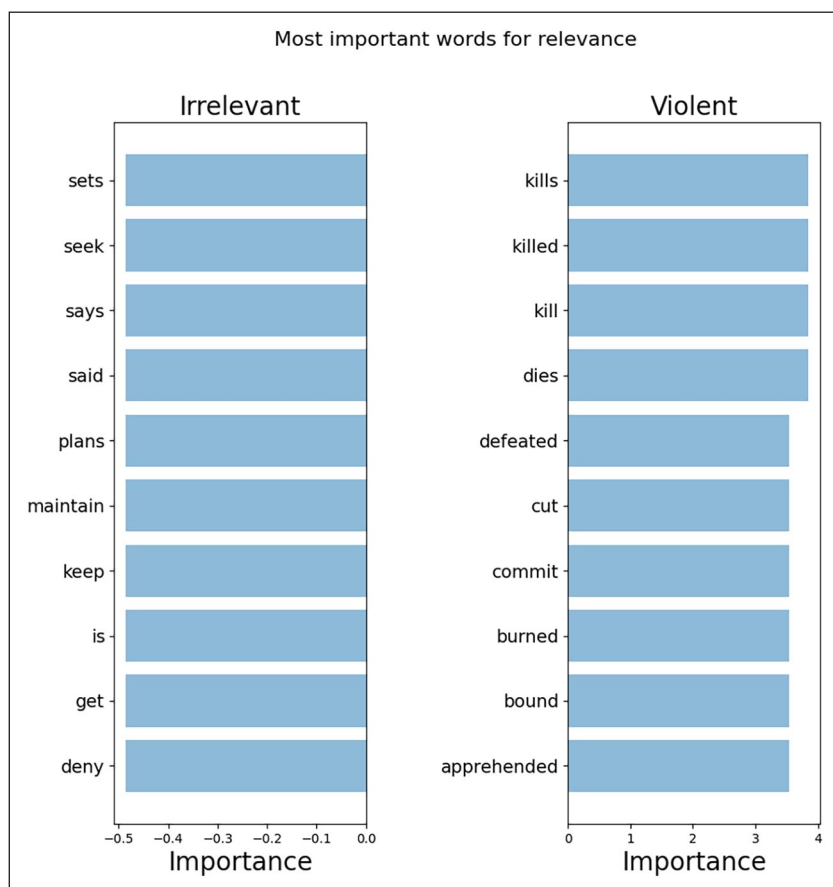


Figure 1. Violent words of top ten importance identified by my logistic regression model.

— a popular, high-coverage sentiment classification resource that assigns words three scores measuring their positivity, negativity, and objectivity (which ‘indicates the presence of opinion’) (Esuli and Sebastiani, 2007: 1). In addition, I randomly sampled 500 articles throughout both periods and blindly tagged key words describing the behavior of Palestinian and Israeli groups and individuals as either violent or nonviolent. This classification produced a word bank I use to train a logistic-regression model augmented with TF-IDF (see Figure 1), closely following methods from Ameisen (2019). My model had an accuracy of 90.6 percent and a precision of 80.6 percent. I used my regression model to classify the violent sentiment of words with an objectivity score less than 1. I have posted my complete content analysis code free and open-source on Github.¹

Sample selection

I perform my analysis on an archived version of the concretely-annotated NYT (Ferraro et al., 2014). All articles are pre-filtered based on time period and their relevance to

Palestine and/or Israel. Each sentence of each article is analyzed individually, and only sentences concerning Palestinian or Israeli subjects are included.

Qualitative analysis

To contextualize the quantitative results, I provide additional qualitative analysis on a number of articles. For each period, I select the articles that were flagged as having the highest counts of passive and active voice for Palestinian and Israeli subjects. First, I analyze the article with the highest number of active voice references with respect to Israeli subjects as well as passive voice references with respect to Palestinian subjects (i.e. these two examples would take the grammatical form ‘Palestinians kill Israelis’ and ‘Palestinians are killed’, respectively). Second, I analyze the article with the highest number of active voice references with respect to Palestinian subjects as well as passive voice references with respect to Israeli subjects (i.e. these two examples would take the grammatical form ‘Israelis kill Palestinians’ and ‘Israelis are killed’, respectively). I base the qualitative analysis mainly on linguistic factors, such as diction and framing, but I also consider factors beyond the article itself, such as decontextualization.

Methodological contribution

Much natural language processing research tries to identify bias based on statistical patterns indiscernible to human readers. In this type of research, the computer *invents* an approximation for how bias is detected based on human examples. Despite opposition from some enthusiasts, most artificial intelligence researchers would agree that computers are *far* from being able to replicate human decision-making. As a result, computer-generated metrics introduce layers of untraceable biases – both human and inhuman in nature. The methods in this study are atypical of much natural language processing research. This study attempts to apply theoretical contributions from previous qualitative work in a quantitative context. Instead of letting the computer infer what bias means, I sublate well-studied human-identified patterns (e.g. indicators based on the passive/active voice). In the section below, I discuss several motivations for increasing human participation in the algorithmic design.

Above all, this study attempts to make a methodological contribution. While the methodology pursued in this research presents more difficult design questions, it discourages the aimless application of ‘flashy’ computational techniques and hopefully can encourage a careful dialogue between quantitative and qualitative analysis.

Limitations of quantitative analysis

While content analysis can be useful, it is never meant to be a substitute for qualitative research. Quantitative analysis can provide breadth that is impossible to achieve qualitatively over the same time scale, but it ultimately sacrifices depth and risks reductionism and overgeneralization. Quantitative research necessarily excludes context. Experts in content analysis expound that the results of any quantitative analysis must be validated with statistics and qualitative measures since they can otherwise be misleading (Grimmer

and Stewart, 2013: 4–5). Some studies have noted bias in content analysis studies of Palestine and Israel since selective methods and unsound analyses can be used to manipulate results in anyone's favor (Kressel, 1987). To account for this, I take three precautions. First, I use pre-trained, well-tested natural language processing models for the majority of my NLP pipeline (NLTK 3.6.2, nd; spaCy, nd). Second, I limit bias during my analysis by using a manually-constructed, data-driven word bank from a representative set of blindly-tagged key words. Finally, I follow up my quantitative analysis with a qualitative validation step, analyzing biased articles in each period by hand.

It is important to contextualize that pre-trained natural language processing models do not provide unbiased metrics. Since they are trained on a corpus of texts that have Orientalist sentiments, the classifiers themselves carry this same anti-Arab, anti-Muslim bias. Studies have found persistent sexist, racist, and Islamophobic bias in highly-used natural language processing toolkits (Abid et al., 2021; Bolukbasi et al., 2016; Bordia and Bowman, 2019; Lu et al., 2020; Nadeem et al., 2020; Shearer et al., 2019; Sheng et al., 2019). I attempt to reduce the influence of ingrained bias in available models in my study. First, I used pre-trained natural language processing toolkits mainly for syntactical parsing (i.e. finding sentence structure). Second, I employ pre-trained sentiment classifiers on individual words only. I do not reveal the subject to the classifier. Third, I train my own model to classify references to violence instead of using pre-trained models. However, these steps alone cannot completely mitigate bias in these models. As a result, all the following content analysis results should be considered in the context of anti-Palestinian bias in artificial intelligence tools, indicating anti-Palestinian bias is likely deeper than these results are capable of measuring.

The First Intifada

Historical context

The First Intifada began in December 1987 and lasted until September 1993. The movement was largely organized by first-time activists enraged by the Israeli occupation of the West Bank and Gaza for the past 20 years (Beinin and Hajar, 2014: 8–9). The Israeli government responded with excessive force to the popular movement, killing '1,422 Palestinians, almost one every day . . . of them, 294, or over 20 percent were minors sixteen and under' (Khalidi, 2021: 172). In comparison, 'one hundred and seventy-five Israelis, 86 of them security personnel, were killed by Palestinians during the same period'. Most US news coverage (including the NYT) did not reflect the true nature of the movement and its human toll (Khalidi, 2021: 172).

Possible sources of bias

During the First Intifada, the NYT staff in control of reporting on Palestine and Israel generally appeared to have pro-Israel biases. Some of the most frequent contributors on articles relating to Palestine and Israel during the First Intifada were AM Rosenthal and Thomas L Friedman. Mearsheimer and Walt (2007: 170) call AM Rosenthal one of many 'passionate defenders of Israel'. In addition, Rosenthal had a reputation for inaccurate

reporting after his botched reporting of the Genovese incident (Levy, 2014). Chomsky (2001) criticizes flaws in Friedman's reporting, including his obsessive focus on analyzing the effects of suicide bombings by Palestinian groups – despite any evidence of the events in their aftermath – and ignorance of Israeli terrorism against Palestinians (pp. 107–108). Mearsheimer and Walt (2007) also note that Friedman 'almost never takes the Palestinians' side or advocates that the United States distance itself from Israel' (p. 170).

During the entirety of the First Intifada, Max Frankel served as Executive Editor of *The New York Times* and had final editing privileges on all articles published. Frankel has even admitted his own pro-Israel bias in his memoir:

I was much more deeply devoted to Israel than I dared to assert . . . Fortified by my knowledge of Israel and my friendships there, I myself wrote most of our Middle East commentaries. As more Arab than Jewish readers recognized, I wrote them from a pro-Israel perspective. (Frankel, 1999: 401–403, cited in Mearsheimer and Walt, 2007: 172)

There is a glaring pro-Israel bias in those reporting on Palestine and Israel.

Content analysis results

The NYT published over 16,000 articles related to Palestine and/or Israel during the First Intifada. Of these, around 40 percent reference Palestinian groups or individuals, while around 93 percent reference Israeli groups or individuals. According to the results of the logistic regression, 11.9 percent of all references to Palestinians used violent language, while only 5.9 percent of all references to Israelis used violent language. Figure 2 shows how the percentage of references that used violent language in the NYT evolved over the First Intifada. At the beginning of the period, violent language was used significantly more in references to Palestinians than in references to Israelis. These statistics became slightly closer towards the end of the Intifada. It is important to note that Israeli violence was actually significantly higher than Palestinian violence during this period (Beinin and Hajar, 2014; Khalidi, 2021), so equality of metric does not mean equality of representation in this case.

The NYT reports referred to Palestinians in the passive voice 15.7 percent of the time. On the other hand, reports referred to Israelis using the passive voice only 6.4 percent of the time. This means the NYT referred to Palestinians in the passive voice more than twice as often as they did Israelis. Figure 3 shows the tone and objectivity scores of these passive voice references. In addition to being referenced more frequently in the passive voice, references to Palestinians were also generally more negative in tone, and references to Israelis in the passive voice were generally more objective. (For example, 'Palestinian killed as clashes erupt with troops' from a 1990 article by Chartrand vs 'Palestinians slay 2 Israeli hikers' from a 1993 article by Greenberg). The use of passive voice de-emphasizes or hides those perpetrating such negative action on Palestinians; this has the rhetorical effect of minimizing the responsibility of Israeli aggressors in causing Palestinian suffering.

Qualitative review of content analysis results

A deep dive into articles highlighted by my content analysis confirms the identified bias and reveals additional biases that take less quantifiable forms. An October 1988

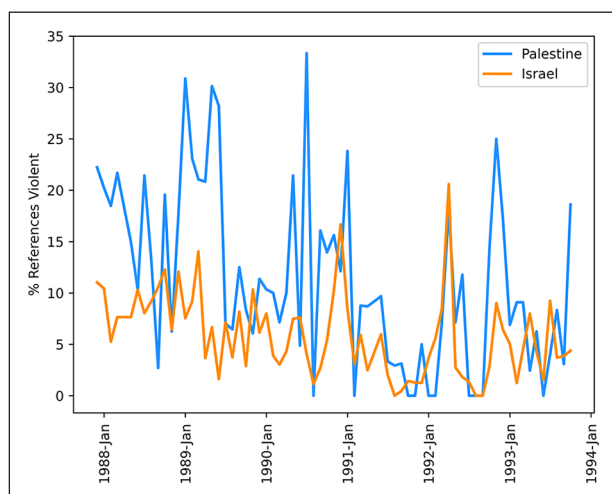


Figure 2. Percent of references to Palestinians and Israelis that used violent language during the First Intifada.

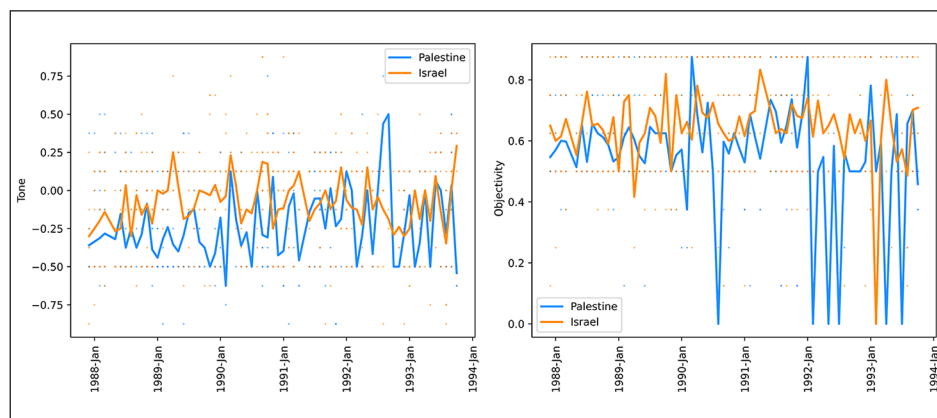


Figure 3. Tone (left) and objectivity (right) of passive voice references to Palestinians and Israelis during the First Intifada. Positive values indicate positive tone, negative values indicate negative tone, and zero values indicate neutrality. More positive objectivity values indicate higher objectivity (i.e. less presence of opinion) in the reference.

article – ‘Proposals for Peace’ – by Thomas L Friedman contained the highest uses of the active voice with respect to Israeli subjects and the passive voice with respect to Palestinian subjects in the First Intifada. In his article, Friedman (1988) begins by introducing a Jewish grocer Sasson – who he believes ‘is the key to a Palestinian–Israeli peace settlement’. Sasson has a clear negative bias against Palestinians, describing them as unwilling to ‘accept a Jewish state’ and implying Palestinians would violently attack Israelis given any chance. Despite Sasson’s violent claims, Friedman proceeds to highlight Sasson’s perspective throughout the article.

Meanwhile, Friedman talks of the Palestinian characters in his news story much less sympathetically. Most of the time, he, or Israeli individuals he quotes in his article, refer to Palestinians in the passive voice. For example, when proposing solutions to the conflict, he quotes an Israeli leader who says ‘only after Palestinians’ rights . . . have been recognized will they be able to focus on their interests’ and ‘the Palestinian government would be granted autonomy’ in his plan. Neither quote specifies the actor who has the responsibility for recognizing the Palestinians’ rights or allowing their government autonomy.

On the other side of the spectrum, a December 1990 article contained the highest uses of the active voice with respect to Palestinians and the passive voice with respect to Israelis in the First Intifada. Despite a high use of the active voice to refer to Palestinians, a qualitative analysis still reveals strong anti-Palestinian sentiment. This qualitative analysis shows that anti-Palestinian bias goes far beyond my quantitative results.

The article, titled ‘Divided Loyalties’ by Joel Brinkley (1990), frames Palestinians living in Jordan as violent and depraved. Brinkley makes overarching, inconsistent claims about the beliefs of all Palestinians, such as ‘Palestinians seem intent on dragging Jordan into political and economic ruin’, and ‘Palestinians are not only openly hostile to the United States, they also want a war.’ One common form of Orientalist bias is to erode the humanity and trustworthiness of the seemingly innocent Arab. In his article, Brinkley describes Yousef Hamden, a Palestinian lawyer as: ‘Dressed in a white shirt and gray silk necktie, hair carefully coiffed, he leans forward in a high-back leather chair behind a rosewood desk. But Palestinian rage burns just below his proper establishment surface.’ Despite his stately appearance, Hamden still ‘burns’ with rage. This description reflects a toxic narrative that no Palestinian, no matter how kind or innocent they appear, can be trusted.

This qualitative analysis not only confirms the bias recorded by my content analysis metrics, but also demonstrates that even articles that do not employ passive voice as a major communicator of bias still convey anti-Palestinian sentiment. My content analysis results scratch the surface of a much deeper pattern of anti-Palestinian bias.

Qualitative review of other sources

A survey of other articles from this period is not shy of blatantly biased headlines – from opinion pieces and approved letters to the editor, inflammatory examples instantly reveal themselves: a letter-to-the-editor ‘Israel and Arab neighbors must bend a little: No more Palestine’ claims Palestine has not existed since 1948 (Neustadter, 1991); an opinion piece ‘Israel has controlled little of Palestine’ claims Israel has only ever controlled 22 percent of Palestinian land (Klein, 1993); and an article written by an NYU professor ‘Retaliation won’t deter Islamic terrorists’ claims young radicalized Muslim men are incorrigible (Majzlin, 1994). The arguments in these articles are quickly disproved with a quick cross-reference to an historical source (Beinin and Hajar, 2014; Khalidi, 2021).

Additionally, articles with Orientalist content hide behind less conspicuous headlines. Throughout the First Intifada, a number of NYT reports were marked by a disproportionate focus on seemingly minor political events in Palestine and Israel. These reports highlighted petty disputes between Palestinian groups or brought up contradiction in Palestinian leaders’ logic and strategy in order to frame Palestinians as disorganized and

illogical (Hijazi, 1987, 1988; *The New York Times*, 1989). This is contrary to the true nature of the Intifada which, as Khalidi (2021: 174) describes, ‘was driven by a broad strategic vision and a unified leadership, and it did not exacerbate internal Palestinian divisions’. For example, a 1995 article titled ‘20 jailed Arab women say no to Israeli offer of freedom’ is particularly inflammatory (Schmemmann, 1995). As is eventually detailed in the article, readers learn that these 20 jailed women refused freedom as part of a coordinated protest of Israel’s imprisonment of other Palestinian women. But the headline immediately throws doubt on the Palestinian cause, propagating a toxic narrative that – even when offered something as undeniable as freedom – Palestinians are bound to respond irrationally.

The Second Intifada

Historical context

Worsening conditions in the Occupied Territories and dissatisfaction with the Oslo ‘peace’ accords sparked the Second Intifada in September 2000, which lasted until February 2005. The Second Intifada was significantly more deadly than the First and saw an increase in Palestinian violence (Beinin and Hajar, 2014). However, Israeli violence far exceeded the extent of Palestinian violence and, in fact, far exceeded the extent of the First Intifada. The Second Intifada left 4,916 Palestinians and 1,100 Israelis dead. Khalidi (2021: 213) explains, ‘Israel turned to heavy weapons, including helicopters, tanks, and artillery, producing even higher Palestinian casualties.’ Despite this disparity in death rate, Palestinians were still scapegoated in US news (pp. 215–216).

Possible sources of bias

Many of the same reporters who covered the First Intifada (e.g. Serge Schmemmann, Joel Greenberg) were frequent contributors to coverage during the Second Intifada as well. However, some new reporters joined the NYT staff, such as popular opinion columnist David Brooks. ‘Today, David Brooks consistently defends Israel’s position’, claim Mearsheimer and Walt (2007: 170) in *The Israel Lobby*. David Brooks is not the only NYT reporter Mearsheimer and Walt criticize during the Second Intifada. In fact, they claim, ‘No one in the *Times*’s stable of regular columnists is a consistent defender of the Palestinians.’ Despite this, some progress in representation was made; during the Second Intifada, Marwan Bishara – a Palestinian journalist – regularly guest-wrote for the NYT (Al Jazeera, 2021). Palestinian representation, however, does not necessarily imply representation of pro-Palestine opinions.

Content analysis results

The NYT published over 17,000 articles related to Palestine and/or Israel during the Second Intifada. Of these, around 49 percent reference Palestinian groups or individuals, while around 93 percent reference Israeli groups or individuals. According to the results of the logistic regression, 15.9 percent of all references to Palestinians used violent

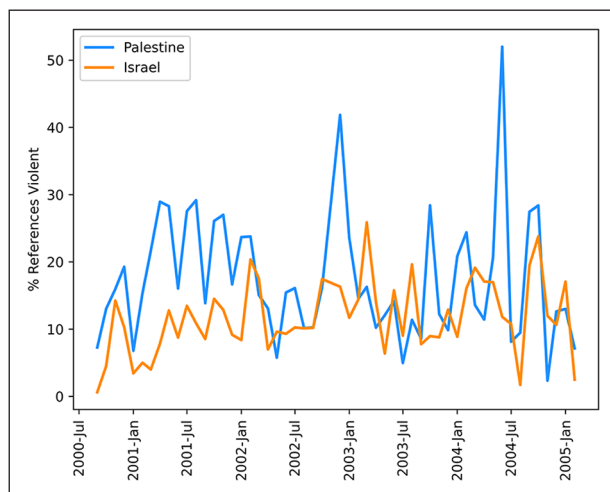


Figure 4. Percentage of references to Palestinians and Israelis that used violent language during the Second Intifada.

language, while 11.7 percent of all references to Israelis used violent language. Figure 4 shows how the percentage of references that used violent language in the NYT evolved over the Second Intifada. The NYT referenced Palestinians using violent language far more often than they did Israelis throughout the entire period.

While both percentages decreased from the First Intifada, Palestinians were still referred to using the passive voice twice as often as Israelis (12.6% and 6.3%, respectively). References to Israelis in the passive voice had more positive tone and were generally seen to be more objective (see Figure 5).

Qualitative review of content analysis results

A qualitative analysis of an article with a high bias score from my content analysis results confirms its Orientalist tone. The article by Douglas Frantz and James Risen (2002), 'A secret Iran–Arafat connection is seen fueling the Mideast fire', has the highest uses of the active voice for Israelis and the passive voice for Palestinians in the Second Intifada. In the article, Frantz and Risen postulate about a weapons exchange between Iran and Palestine. However, they rarely directly cite Palestinian individuals. They mainly refer to Palestinian groups in the passive voice and often indiscriminately refer to 'Palestinian terrorists'. In fact, Frantz and Risen both start and end their article with endorsements of Israeli officials' accusations – thus framing the entire article from a pro-Israeli perspective.

While the bias in Frantz and Risen's article is clear from both quantitative and qualitative results, even articles that use a more active voice to describe Palestinians and a more passive voice to describe Israelis contain anti-Palestinian bias. James Bennet's (2004) article, 'Isolated and angry, Gaza battles itself, too', has the highest uses of active voice

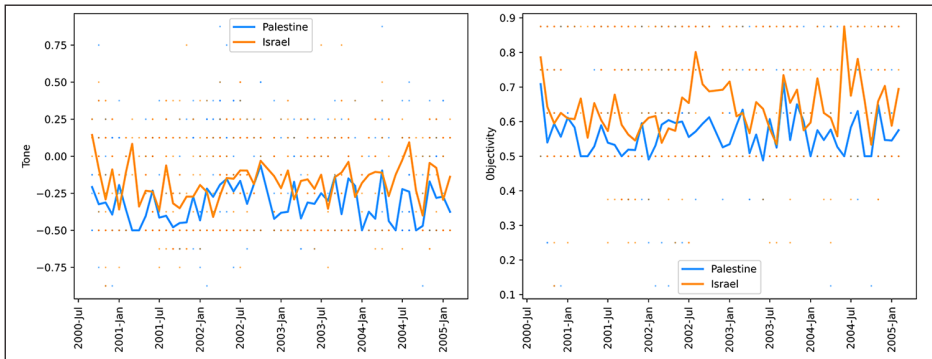


Figure 5. Tone (left) and objectivity (right) of passive voice references to Palestinians and Israelis during the Second Intifada. Positive values indicate positive tone, negative values indicate negative tone, and zero values indicate neutrality. More positive objectivity values indicate higher objectivity (i.e. less presence of opinion) in the reference.

for Palestinians and passive voice for Israelis in the Second Intifada. However, Bennet employs the active voice to highlight disorganization between Palestinian groups and misrepresent Palestinian opinion. Bennet's article disproportionately focuses on Israeli-favored political candidates who could rise to power in Gaza after Israel began its withdrawal – for instance, Muhammad Dahlan, who is 'favored by Israeli, European and American officials as strong enough to run Gaza'. It is clear he hides the opinions of those who dissent with Dahlan, who speaks from a pro-Israeli perspective. Consequently, Bennet's report presents only one perspective – a pro-Israeli one.

Qualitative review of other sources

Throughout the Second Intifada, bias persisted in the NYT coverage, masquerading behind more innocent headlines. Two examples in particular stand out from this period. Brooks's (2005) Opinion piece 'What Palestinians?' provides a grossly oversimplified analysis. After arguing that Israel has begun to 'disengage' from Palestinian affairs without much more than anecdotal evidence, Brooks (2005) comments, 'The Palestinians richly deserve to be left behind. Even now they expect Israel to allow Palestinian trucks to cross its border, even though both sides know some significant portion will contain bombs designed to kill Jews.' Brooks' brazen, unsubstantiated accusation that a 'significant portion' of Palestinians crossing the border (which is in fact monitored by the Israeli army) want to murder Jews quickly renders his opinion piece no more than a racist rant.

A devolving perspective of Palestine

Over the First and Second Intifadas, the content analysis results reveal a disproportionate anti-Palestinian sentiment. Between the periods, the nature of the results remains similar with a slight increase in positive representation of Palestinians in the Second Intifada.

However, in both periods – of all articles relating to Palestine and/or Israel – less than 50 percent referenced Palestine while over 90 percent of the articles referenced Israel. The Israeli narrative dominates the discourse in both periods. In addition, during both periods, references to Palestinians were more than 1.5 times as likely to use violent language than references to Israelis (about 2 times in the First Intifada and about 1.5 times in the Second). Finally, Palestinians were consistently referred to more frequently in the passive voice than Israelis (about 2.5 times more than Israelis in the First Intifada and about 2 times more than Israelis in the Second Intifada).

The qualitative analysis of the articles additionally confirmed biased content during both periods. In addition, my qualitative analyses began to explore how anti-Palestinian bias in the NYT reaches even further than my content analysis results. While my analyses revealed a number of telling factors, there are many Orientalist techniques – such as decontextualization and selective reporting – that even state-of-the-art content analysis methods cannot yet identify.

While the content analysis may show a slight improvement in representation of Palestinians, considering the historical context of the periods negates any positive progress. Although casualty statistics do not provide a wholly reflective perspective, the difference in casualty numbers between the periods is striking. Overall casualties (both Palestinian and Israeli) during the Second Intifada exceeded the First Intifada by over 4.6 times (Khalidi, 2021: 172, 213). Despite this increase, there were still hundreds more Palestinian casualties during the First Intifada and thousands more Palestinian casualties during the Second Intifada than Israeli casualties during the Second Intifada (Khalidi, 2021: 172, 213). The magnitude of anti-Palestinian violence increased significantly during the Second Intifada. However, the NYT reporting of this violence did not increase proportionally. The stagnation of the content results contrasted with the stark increase in violence between the Intifadas suggests that the NYT reporting became more biased against Palestinians during the Second Intifada.

Consequently, these content analysis results support the theory that US news coverage painted a more negative picture of Palestinians during the Second Intifada than the First, unfortunately at a time when Palestinians may have needed international support the most.

Conclusion

This study provides a broad overview of the evolution of bias in the NYT coverage of Palestinians and Israelis during the First and Second Intifadas. The content analysis revealed clear patterns of bias against Palestinians in the NYT coverage through two main linguistic features: (1) a disproportionate use of the passive voice to refer to negative or violent action perpetrated towards Palestinians, and (2) use of more negative and violent rhetoric in reference to Palestinians compared to Israelis. In conjunction with the historical context of the increased violence towards Palestinians in the Second Intifada, the minimal change in content analysis results between the two periods implies an increase in overall anti-Palestinian sentiment in the NYT news coverage.

The NYT is a highly influential newspaper to many Americans – from civilians to government officials – and the sentiment of US news coverage towards Palestine is truly

a life-or-death matter since the US is such a powerful foreign entity in the region. This newspaper serves as a case study of how Orientalist bias is pernicious in US news coverage, and is more far reaching than many may anticipate.

In a world of mis- and disinformation, we rely on whistleblowers to pinpoint and denounce media bias in an industry that prides itself on its ‘objectivity’. This qualitative investigative work – which can take the form of an academic paper to a social media post – is built (whether knowingly or unknowingly) on well-verified metrics from the social sciences. However, as the volume of media constantly expands through digital platforms, it becomes more and more difficult for individuals and activist groups to keep up with the quantity of biased material. Denouncements of bias are often drowned out in ‘both-sides’ narratives and moral relativism. Careful computational analysis, like the methodology in this study, can help us gather evidence at a rate that matches mass media production – building on a strong and underappreciated body of qualitative work that establishes an historical precedent for anti-Palestinian bias in US news. In an era where quantitative and qualitative analysis are often pitted against each other, this work shows how well-studied metrics from the social sciences can be harnessed by large-scale computational research. These automated tools can support global documentation efforts of anti-Palestinian bias.

Acknowledgements

This study is dedicated to the people of Palestine and their struggle for liberation. This work would not have been possible without the underappreciated and unrewarded (often punished) labor of countless Palestinians, who have historically called out the NYT and other media outlets that misrepresent their struggle. The author gratefully acknowledges Professor Richard Samuels, who provided insightful feedback throughout the development of this study, and Professor Pouya Alimagham, who has provided invaluable mentorship and feedback on the research. She is also grateful to Professor Richard Nielsen, who provided a number of content analysis resources. The author extends warm appreciation to the anonymous reviewers of this article, whose careful and critical feedback greatly improved the study. A version of this article was published online as a pre-print before submission to *Media, War & Conflict* (accessible at *The New York Times Distorts the Palestinian Struggle* (mit.edu)).

Funding

The author received no financial support for the research, authorship, and publication of this article.

ORCID iD

Holly M Jackson:  <https://orcid.org/0000-0001-6898-6767>

Data availability

All supporting code can be accessed on Github at: https://github.com/hollyjackson/NYT_Content_Analysis.

Note

1. https://github.com/hollyjackson/NYT_Content_Analysis

References

- Abid A, Farooqi M and Zou J (2021) Persistent anti-Muslim bias in large language models. *arXiv* pre-print: arXiv:2101.05783.
- Abunimah A (2002) For NPR, violence is calm if it's violence against Palestinians. Report, Fairness & Accuracy in Reporting, 10 January. Available at: <https://web.archive.org/web/20080514145307/http://www.fair.org/activism/npr-israel-quiet.html> (accessed 29 May 2023).
- Ackerman S (2001) Al-Aqsa intifada and the U.S. media. *Journal of Palestine Studies* 30(2): 61–74.
- Ackerman S (2009) FAIR challenges CBC Ombud's Report. Report, Fairness & Accuracy In Reporting, 4 February. Available at: <https://fair.org/press-release/fair-challenges-cbc-ombuds-report/> (accessed 29 May 2023).
- Al Jazeera (2021) Marwan Bishara: Al Jazeera news, 29 April.
- Al-Sarraj WF and Lubbad HM (2018) Bias detection of Palestinian/Israeli conflict in western media: A sentiment analysis experimental study. In: *International Conference on Promising Electronic Technologies (ICPET)*, Deir El-Balah, Palestine, 98–103. New York, NY: IEEE.
- Altheide DL (1984) Media hegemony: A failure of perspective. *Public Opinion Quarterly* 48(2): 476–490.
- Altschull HJ (1995) *Agents of Power: The Media and Public Policy*. London: Pearson.
- Ameisen E (2019) How to solve 90% of NLP problems: A step-by-step guide. *KDnuggets*. Available at: <https://www.kdnuggets.com/2019/01/solve-90-nlp-problems-step-by-step-guide.html> (accessed 10 May 2021).
- Aziz M (2007) 'An analysis of print media coverage of the Palestinian–Israeli conflict during the second Israeli invasion of Lebanon in 2006', MA thesis, Clemson University, Clemson, SC.
- Bazian H (2015) The islamophobia industry and the demonization of Palestine: Implications for American studies. *American Quarterly* 67(4): 1057–1066.
- Beinin J and Hajar L (2014) Palestine, Israel and the Arab-Israeli conflict: A primer. Report, Middle East Research and Information Project, Washington, DC.
- Bennet J (2004) Isolated and angry, Gaza battles itself, too. *The New York Times*, 16 July.
- Bolukbasi T et al. (2016) Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: *30th Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain, 5–10 December 2016, 4356–4364. Red Hook, NY: Curran Associates, Inc.
- Bordia S and Bowman SR (2019) Identifying and reducing gender bias in word-level language models. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*, Minneapolis, MN, 3–5 June, 7–15. Stroudsburg, PA: Association for Computational Linguistics.
- Boudana S (2011) A definition of journalistic objectivity as a performance. *Media, Culture & Society* 33(3): 385–398.
- Brinkley J (1990) Divided loyalties. *The New York Times*, 16 December.
- Brooks D (2005) What Palestinians? *The New York Times*, 17 November.
- Caballero JX (2010) 'The impact of media bias on coverage of catastrophic events: Case study from The New York Times' coverage of the Palestine/Israel Conflict', undergraduate thesis, University of Pittsburgh, Pittsburgh, PA.
- Chartrand S (1990) MIDEAST TENSIONS: Palestinian killed as clashes erupt with troops. *The New York Times*, 4 November.
- Chomsky N (2001) Middle East terrorism and the American ideological system. In: Said E and Hitchens C (eds) *Blaming the Victims: Spurious Scholarship and the Palestinian Question*. London: Verso, 97–147.

- Cohen, R (2021) An Israeli Death and the Tangled Conflict Left Behind. *The New York Times*, 28 May.
- Esuli A and Sebastiani F (2007) SentiWordNet: A high-coverage lexical resource for opinion mining. Technical Report, Istituto di Scienza e Tecnologie dell'Informazione, Pisa, Italy.
- FAIR (2002) In U.S. media, Palestinians attack, Israel retaliates. Report, Fairness & in Reporting, 4 April. Available at: <https://web.archive.org/web/20090115042817/http://www.fair.org/activism/network-retaliation.html> (accessed 29 May 2023).
- Ferraro F et al. (2014) Concretely annotated corpora. In: *Proceedings of the NIPS Workshop on Automated Knowledge Base Construction (AKBC)*, Montreal, Canada, 13 December.
- Frankel M (1999) *The Times of My Life and My Life with The Times*. New York, NY: Random House.
- Frantz D and Risen J (2002) A NATION CHALLENGED: TERRORISM: A secret Iran–Arafat connection is seen fueling the mideast fire. *The New York Times*, 24 March.
- Frazer AK and Miller MD (2008) Double standards in sentence structure. *Journal of Language and Social Psychology* 28(1): 62–71.
- Friedman TL (1988) Proposals for peace. *The New York Times*, 30 October.
- Glasser TL (1988) Objectivity precludes responsibility. In: Hiebert RE and Reuss C (eds) *Impact of Mass Media: Current Issues*. White Plains, NY: Longman, 44–51.
- Greenberg J (1993) Palestinians slay 2 Israeli hikers; Israelis kill suspected guerrilla. *The New York Times*, 10 October.
- Grimmer J and Stewart BM (2013) Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21(3): 267–297.
- Hackett RA (1984) Decline of a paradigm? Bias and objectivity in news media studies. *Critical Studies in Mass Communication* 1(3): 229–259.
- Hémanus P (1976) Objectivity in news transmission. *Journal of Communication* 26(4): 102–107.
- Hijazi IA (1987) Rival Palestinians meet in effort to end division. *The New York Times*, 14 April.
- Hijazi IA (1988) Palestinian factions divided on forming a state in exile. *The New York Times*, 11 August.
- Jagota A (2020) Named entity recognition in NLP: Medium, towards data science. Available at: <https://towardsdatascience.com/named-entity-recognition-in-nlp-be09139fa7b8> (accessed 29 May 2023).
- Kazenin KI (2008) The passive voice. In: 2. *Halbband Language Typology and Language Universals 2. Teilband*. Berlin: De Gruyter Mouton, 899–916. DOI: 10.1515/9783110194265-004.
- Keras (nd) Available at: <https://keras.io/> (accessed 10 May 2021).
- Khalidi R (2021) *The Hundred Years' War on Palestine: A History of Settler Colonialism and Resistance, 1917–2017*. New York, NY: Picador.
- Kingsley P (2021) As Gaza war escalates, new front opens in Israeli cities. *The New York Times*, 12 May.
- Klein M (1993) Israel has controlled little of Palestine. *The New York Times*, 11 September.
- Kressel NJ (1987) Biased judgments of media bias: A case study of the Arab–Israeli dispute. *Political Psychology* 8(2): 211–227.
- Levy N (2014) How A.M. Rosenthal made a crime into a crisis. *Politico*, 5 March.
- Lu K et al. (2020) Gender bias in neural natural language processing. In: Nigam V et al. (eds) *Logic, Language, and Security: Essays Dedicated to Andre Scedrov on the Occasion of His 65th Birthday*. New York, NY: Springer International Publishing, 189–202.
- McQuail D (1992) *Media Performance: Mass Communication and the Public Interest*. London: Sage.
- McQuail D (1994) *Mass Communication Theory: An Introduction*. London: Sage.

- Majzlin L (1994) Retaliation won't deter Islamic terrorists; Palestine's boundaries. *The New York Times*, 27 October.
- Mearsheimer JJ and Walt SM (2007) *The Israel Lobby and U.S. Foreign Policy*. New York, NY: Farrar, Straus and Giroux.
- Nadeem M, Bethke A and Reddy S (2020) StereoSet: Measuring stereotypical bias in pre-trained language models. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, online, 1–6 August 2021, 5356–5371. Stroudsburg, PA: Association for Computational Linguistics.
- NLTK 3.6.2 (nd) Available at: <https://www.nltk.org/> (accessed 10 May 2021).
- Neustadter P (1991) Israel and Arab neighbors must bend a little: No more Palestine. *The New York Times*, 21 June.
- The New York Times* (1989) P.L.O. rejects warning by U.S. on attacks. *The New York Times*, 5 March.
- The New York Times* (nd) Ethical journalism: A handbook of values and practices for the news and opinion departments. Available at: <https://www.nytimes.com/editorial-standards/ethical-journalism.html#> (accessed 1 March 2023). NLTK 3.6.2 (nd) Available at: <https://www.nltk.org/> (accessed 10 May 2021).
- Roy S (2010) Reflections on the Israeli–Palestinian conflict in U.S. public discourse: Legitimizing dissent. *Journal of Palestine Studies* 39(2): 23–38.
- Said E (2001) Introduction. In: Said E and Hitchens C (eds) *Blaming the Victims: Spurious Scholarship and the Palestinian Question*. London: Verso, 1–19.
- Schmemmann S (1995) 20 jailed Arab women say no to Israeli offer of freedom. *The New York Times*, 9 October.
- Schudson M (2001) The objectivity norm in American journalism. *Journalism* 2(2): 149–170.
- scikit-learn (nd) Available at: <https://scikit-learn.org/stable/> (accessed 10 May 2021).
- Shearer E et al. (2019) *Racial Bias in Natural Language Processing*. Report, Oxford Insights, Malvern.
- Sheng E et al. (2019) The woman worked as a babysitter: On biases in language generation. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, Hong Kong, China, 3–7 November 2019, 3407–3412. Stroudsburg, PA: Association for Computational Linguistics.
- Siddiqui U and Zaheer OA (2018) *50 Years of Occupation: A Sentiment and N-Gram Analysis of U.S. Mainstream Media Coverage of the Israeli Occupation of Palestine*. Report, 416Labs, Toronto, ON.
- Sirhan N (2021) *Reporting Palestine–Israel in British Newspapers: An Analysis of British Newspapers*. London: Palgrave Macmillan.
- Sögüt S (2018) Ideology in the news through active, passive sentences and nominalization: A study on the terrorist attack in Ankara reported in British and American newspapers. *Journal of Language and Linguistic Studies* 14(1): 162–177.
- spaCy (nd) Available at: <https://spacy.io/> (accessed 10 May 2021).
- Tuchman G (1972) Objectivity as strategic ritual: An examination of newsmen's notions of objectivity. *American Journal of Sociology* 77(4): 660–679.
- Weir A (2005) Off the charts: Accuracy in reporting of Israel/Palestine, *The New York Times*. Report, *If Americas Knew*, Vancouver, WA.
- Zelizer B, Park D and Gudelunas D (2002) How bias shapes the news: Challenging the *New York Times'* status as a newspaper of record on the Middle East. *Journalism* 3(3): 283–307.

Author biography

Holly M Jackson is a human rights studies Master's student at Columbia University concentrating in imperialism and settler-colonialism in the Middle East. She holds a Bachelor's in electrical engineering and computer science from the Massachusetts Institute of Technology. Starting fall 2023, she will begin a PhD in computer science at UC Berkeley. She works on interdisciplinary applications of computer science, from astrophysics to history to political problems. More information is available at her website: holly-jackson.com.