Gender bias, social bias, and representation: 70 years of Bollywood

Graphical abstract

Highlights

- Applies cutting-edge natural language processing methods to movie subtitles
- Considers 700 popular Bollywood movies spanning 70 years to analyze evolving trends
- Tracks evolving biases toward gender, color, geographic region, and religion
- Compares movies from Bollywood, Hollywood, and critically acclaimed world movies

Authors

Kunal Khadilkar, Ashiqur R. KhudaBukhsh, Tom M. Mitchell

Correspondence

axkvse@rit.edu

In brief

We present a comprehensive analysis of gender and social biases present in English subtitles of popular Bollywood movies, spanning 70 years, using cutting-edge natural language processing techniques. We also consider popular Hollywood movies and movies nominated for the Academy Awards to compare and contrast our findings. Our study uncovers social biases along the lines of colorism, son preference, and geographic and religious representation.
Gender bias, social bias, and representation: 70 years of Bollywood

Kunal Khadilkar,1,4 Ashiqur R. KhudaBukhsh,2,4,5,* and Tom M. Mitchell3

1School of Computer Science, Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
2Golisano College of Computing and Information Sciences, Software Engineering Department, Rochester Institute of Technology, 20 Lomb Memorial Drive, New York, NY 14623, USA
3School of Computer Science, Machine Learning Department, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
4These authors contributed equally
5Lead contact
*Correspondence: axkvse@rit.edu
https://doi.org/10.1016/j.patter.2021.100409

SUMMARY

We use a suite of cutting-edge natural language processing methods to quantify and characterize societal and gender biases in popular movie content. Our data set consists of English subtitles of popular movies from Bollywood—the Mumbai film industry—spanning 7 decades (700 movies). In addition, we include movies from Hollywood and movies nominated for the Academy Awards for contrastive purposes. Our findings indicate that while the overall portrayal of women has improved over time in popular movie dialogues from both Bollywood and Hollywood, modern films still exhibit considerable gender bias and are yet to achieve equal representation among genders. We also observe a strong bias favoring fair skin color in Bollywood content that occurred consistently across all time periods we considered. While our geographic representation analysis indicates improved inclusion over time for several Indian states, it also reveals a long-standing under-representation of many northeastern Indian states.

INTRODUCTION

What types of social biases can we analyze and detect through the lens of a diachronic corpus of popular entertainment? In this paper, we focus on Bollywood, also known as the Mumbai film industry, and analyze a curated corpus of film subtitles for the last 70 years. While Bollywood is an entertainment industry worth billions and has a target audience of 1.2 billion people, little or no work exists that has analyzed a wide range of social biases and signals that can be uncovered through a systematic study of these popular films spanning decades. In this work, we contrast our findings with an analogous corpus of Hollywood films, and for a specific subset of research questions, we extend our analysis to world movies.

Our primary focus in this work is gender attitudes and bias (a preliminary version of the work detailed in this article, containing...
In this paper, we analyze a broad set of research questions (described in the following section) using a suite of cutting-edge NLP methods. While many of our explored research questions have received prominence in prior social science literature, we offer a scale unmatched in previous studies. For instance, Rao analyzes the portrayal of women in 19 Bollywood films, but our analysis considers 700 films spanning 70 years for the same research question. Our quantitative comparative approach, contrasting Bollywood with Hollywood, is also new to social science research.

In our mixed method analyses, we identify that (1) some of the gender biases observed in Bollywood are very much present in its Western counterpart; (2) a positive trend is witnessed in observing reduced biases with progress of time; and (3) a similar trend is observed in religious and geographic representation, with a considerable scope for improved diversity and inclusion.

Research questions and paper road map
In this paper, we broadly focus on four research questions that examine (1) gender attitudes and biases, (2) attitudes and biases toward geographic regions, (3) attitudes toward religions and religious representations, and (4) information about the economy and national priorities. In what follows, we show the motivation of each of our research questions and then present a road map to the rest of the paper.

Gender attitudes and biases
A key focus of our paper is portrayal of women in popular Bollywood and Hollywood content. Prior studies indicate that even a simple measure of representation such as relative gendered pronoun usage may reveal important trends in the evolving nature of status of women over time. While the economic growth of India has seen a steady rise over the last 30 years, women’s labor force participation has seen a sharp decline during this period.

During 2017–18, the labor force participation rate of women in India was 21.05%, a substantially lower participation rate than the world average of 47.43% (source: World Bank data, https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?end=2019&locations=IN-1W&most_recent_value_desc=true&start=2010). Occupational stereotypes inferred from linguistic signals reveal important insights about aggregate gender attitudes toward certain professions. We thus analyze gendered pronoun usage and historical trends in occupational stereotypes to answer our first research question:

**RQ 1.1:** How is gender bias reflected through movie dialogues in the Bollywood and Hollywood movie industries?

According to the last two decennial censuses conducted in India, the overall sex ratio (computed as the number of women per 1,000 men) has improved from 933 to 940. However, in the 2011 census, the lowest ever child sex ratio (CSR, computed as the number of female children per 1,000 male children in the age group of 0–6 years) of 914 was recorded. Son preference in India is a well-documented phenomenon, and skewed sex ratios, female feticide, and higher child mortality rates for girls have attracted policymakers’ attention. In order to prevent female feticide, in 1994, the Parliament of India enacted the...
Pre-Conception and Pre-Natal Diagnostic Techniques Act also known as the Prohibition of Sex Selection Act that effectively rendered prenatal sex discernment illegal. Beyond existing research questions involving gendered pronoun usage and occupational stereotypes, we are thus interested in examining the evolving trend of son preference in Bollywood content as our second research question:

**RQ 1.2**: Does Bollywood reflect the well-documented son preference in medical and social science research?

Beyond occupational stereotypes and representational research questions, our study delves into a social bias that has received considerable prominence in social science research: association of beauty and fairness of skin in India. Skin color biases have been reported in the context of fairness beauty products, Indian arranged marriages, and surprisingly, political outcomes in India. While Shevde raises an important point that several Bollywood celebrities endorse skin-whitening products, to the best of our knowledge, no large-scale analysis of fair skin color and beauty in Bollywood (or Hollywood) content exists thus far. We seek to address this gap through our next research question:

**RQ 1.3**: Is beauty associated with fair skin in the movie dialogues describing women?

External factors may influence biased gender attitudes. Certain retrograde practices such as dowry can influence son preference as a girl child might be looked upon as a financial burden. The dowry system has plagued Indian society for a long time. Dowry refers to a transaction of tangible financial objects in the form of durable goods, cash, and real or movable property between the bride’s family and the bridegroom, his parents, and his relatives as a condition of the marriage. Although legally dowry has been prohibited in India since 1961, this practice has continued well after its legal prohibition and has a strong link to social crises such as female feticide, domestic abuse and violence, and dowry deaths. However, while the practice continues, recent studies have reported positive changes in society where the general attitude toward the system has become negative. Since this retrograde practice is interlinked with so many crises, the dowry system in India has received attention from the social science research community for decades. In this paper, we analyze the...
sentiment around this practice over the last 70 years through our next research question:  
**RQ 1.4:** How has the sentiment around retrograde social practices such as dowry evolved?  

**Attitudes toward religions and religious representation**  
With six major religions, 22 languages, and 700 dialects, religion and languages are key diversity factors in pluralistic India. The shifting nature of the demographic balance between the two major religions in India, Hinduism and Islam, and its possible interpretations have a long history of use in political debates. The Indian subcontinent has faced two major partitions over the last 70 years, which have resulted in considerable religious turmoil and multiple riots. Analyzing religions in Bollywood films, both in terms of perception and representation, forms our next research question:  
**RQ 2:** How are religions perceived in movies? Can we gain an insight into the religious representation of a country through a film corpus spanning 70 years?  

**Attitudes and biases toward geographic regions**  
From sports participation to linguistic debates, regional politics has been a recurrent theme in Indian political discourse. As of 2021, India has 28 states and eight union territories. While geographic representation with a key focus on under-represented states in northeast India has been studied in print news medium before, no prior work has studied the evolving nature of geographic inclusion in Bollywood content. Our study thus adds a valuable data point to understanding geographic inclusion in popular cultural products through the following research question:  
**RQ 3:** How has geographic representation evolved over time in Bollywood content? Which geographic areas have been consistently under-represented in the Mumbai film industry?  

**Information about the economy and national priorities**  
Finally, we are interested in exploring if broad trends in economy and national priorities can be tracked from popular entertainment data. Recent work has shown language models can be used to aggregate opinions and track evolving national priorities from social media data. However, such work focused on a shorter time horizon and tracked shifting priorities on a month-by-month basis. In this work, we explore if similar techniques can be applied to analyze historical trends spanning multiple decades in our final research question:  
**RQ 4:** Can we extract economic signals through popular film dialogues? Can we track evolving national priorities from popular entertainment?

**Paper road map**  
The rest of the paper is organized as follows. After describing the relevant literature in "related work," we devote one separate section to each of the four research questions we investigate. *Results: gender attitudes and biases* investigates gender attitudes and biases described in research question **RQ 1** (gender attitudes and biases section). *Results: religion* examines religious representation described in research question **RQ 2** (attitudes toward religions and religious representation section) and *results: attitudes and biases toward geographic regions* investigates geographic representation presented in research question **RQ 3** (attitudes and biases toward geographic regions section). Finally, *economic signals and national priorities* presents our findings on research question **RQ 4** (information about the economy and national priorities section). In "discussion," we
summarize the major takeaways of our study and describe some of the limitations of our work and then describe the materials and methods used in our study in “experimental procedures.”

Related work

The NLP literature focusing on gender and societal biases can be loosely categorized into two broad categories: descriptive and prescriptive. The descriptive (bias evaluation) class of methods presents quantitative frameworks to understand, measure, and analyze bias (e.g., bias in word embeddings,33–35 downstream applications,33,36–38 or large-scale corpora39,39). The prescriptive (bias mitigation) class of algorithms aims to debias using a broad range of techniques.40–42 A comprehensive survey can be found in Garrido-Muñoz et al.43 Since our work is directly related to the former, we next present the bias evaluation side of the literature in greater depth.

In the entertainment industry, bias domain, existing lines of work focus on a single Bollywood movie44 or a small subset of movies.45 Madaan et al.46 focused on plot points and film information taken from Wikipedia. We consider a different and potentially richer data set of film subtitles spanning 70 years. We contrast our work with Hollywood and award-winning world movies, and our analyses cover a broader set of aspects such as retrograde social practices, uncovering subtler biases and highlighting geographic and religious under-representation. Unlike previous work on movie subtitles,45 our focus is on Bollywood content largely ignored by the information science research community so far.

Unlike our focus on popular Bollywood movies, studies analyzing gender stereotypes across different languages35 and detecting bias in word embeddings have previously used books and news data sources for their analyses. Thematically, our work is related to other comprehensive analyses of biases present in different data sources such as biases in history text books32 and narrative tropes.33 Gala et al.33 performed a study to uncover highly gendered tropes and the inherent topics trending within them, while Lucy et al.48 presented a comprehensive study of 15 US history text books used in Texas between 2015 and 2017. Similar to our study, Lucy et al.48 employed a wide array of NLP techniques that revealed several biases along the lines of race and gender and also tied the findings with the political composition of the individual counties.

While most of the previous work in the entertainment industry has revolved around gender biases and portrayal of characters in the movies, Sheth et al.47 provided a qualitative analysis to showcase the growing culture and bias for fair skin in the film industry. The authors also presented detailed evidence to indicate how the film industry played a part in emboldening various stereotypes prevalent in India. Mishra48 provides an even more higher-level study on the finer nuances that exists in Indian society when it comes to discrimination based on skin color. A number of media articles, tabloids, and blogs39–41 give examples of objectionable lyrics or dialogues in blockbuster movies.

A preliminary version of the work detailed in this article, containing a small subset of results and evaluated on a smaller data set, appeared in Khadilkar and KhudaBukhsh.1 The current paper extends that previous work in the following key ways. First, our analysis of gender bias (1) includes diachronic word
embedding analysis and word embedding association tests (WEATs), (2) is grounded on well-established lexicons, (3) looks into subtler signals such as son preference, (4) tracks retrograde social practices such as dowry, and (5) considers additional data sources (e.g., world movies). Second, our work tackles important additional questions on geographic and religious representation unaddressed by the previous study. Finally, we look into questions related to economic signals and evolving national priorities not explored in the earlier version.

In this paper, we explore a wide array of NLP techniques to analyze our research questions through the lens of popular movies:

1. simple count-based statistics relying on highly popular lexicons and gender representation studies;
2. cloze test, an analysis technique that has a solid grounding in psycholinguistics literature. To the best of our knowledge, for the first time, we explore a recent technique previously used to mine political insights in the context of uncovering social biases. Through a series of cloze tests on a language model fine-tuned on our data sets, we present our findings;
3. analysis of aligned diachronic word embedding spaces using recently proposed techniques;
4. free form text completion using GPT-2 for a novel task of tracking economic signals.

We employ this broad suite of NLP techniques on a novel domain of popular entertainment. We also present relevant literature when we describe these techniques.

Geographic and community representation in India has been studied by various political and social scientists. Bhargava showcases the population under-representation in political settings, while Chongloi focuses on under-representation of northeast India in mainstream newspapers. Our work complements this research and presents corroborating evidence from a very different data source.

RESULTS

Gender attitudes and biases

As described in the gender attitudes and biases section, our research question examining gender biases (RQ 1) in movie dialogues consists of four sub-parts. In what follows, we investigate each of these research questions in individual subsections.

RQ 1

RQ 1.1: How is gender bias reflected through movie dialogues in the Bollywood and Hollywood movie industries?

We investigate RQ 1.1 using a diverse set of techniques that includes (1) gendered pronoun usage (RQ 1.1 section), (2) WEAT (word embedding associated test section), (3) aligned diachronic word embeddings (diachronic word embeddings section), and (4) cloze tests (cloze tests section).

Gendered pronoun usage

Following extensive literature on gendered pronouns’ relative distributions and their implications, we start with a simple measure of gender representation: relative occurrence of pronouns of each gender (men: he, him; women: she, her). Let denote the number of times a token appears in a corpus. We define male pronoun ratio (MPR) as follows: .

Figure 1 plots MPR of our decade-wise movie data sets and contrasts with the MPR computed using google n-grams. Our results indicate that, even now, both Bollywood and Hollywood exhibit a comparable skew in gendered pronoun usage.

Word Embedding Association Test

A powerful way to operationalize the notion of words being close (or far) from one another is to employ a method that embeds each word as a vector in a high-dimensional space (referred to as an embedding) and using the proximity of any two words in that space as a measure of closeness. First introduced in Caliskan et al., the WEAT is a well-known statistical test analogous to the implicit association test (IAT) for quantifying biases in text data. WEAT computes the difference in relative cosine similarity between two sets of target words (e.g., occupations) and two sets of attribute words (e.g., male gendered pronouns and female gendered pronouns). The test produces a score within the range of $-1$ to $1$. In our case, a WEAT score of 0 indicates no bias, and a positive (negative) score indicates bias toward men (women).

Figure 2 presents the WEAT scores of Bollywood and Hollywood computed for three non-overlapping time periods. We find that (1) the average WEAT scores across both industries reduced over time; and (2) as compared to Bollywood, for any given time period, Hollywood exhibits less gender bias.

How do award winning foreign feature films compare with Bollywood and Hollywood in addressing gender equality? Does genre make any difference? As a follow-up investigation, we compare the WEAT scores of Bollywood and Hollywood with the WEAT score of a set of critically acclaimed world movies nominated in the foreign film category at the Academy Awards. Our results summarized in Figure 3 showcase that the average WEAT score obtained for nominated foreign feature films is the lowest compared with the average WEAT score for Bollywood and Hollywood.

Within Bollywood and Hollywood, we next examine the influence of genres. Adventure/action and romance are the two most popular genres across different industries, with hundreds of films released every year. Action films generally tend to be male-dominated, compared to romantic films, and hence are more likely to be biased toward men. We explore this hypothesis using WEAT for these genres. For a given movie industry and

---

**Table 3. Percentage increase in average valence score for cloze test completions between old movies and new movies**

| Gender   | Old    | Hollywood (%) | New     |
|----------|--------|---------------|---------|
| Women    | 22.84  | 7.55          |         |
| Men      | 6.00   | 15.60         |         |

**Table 4. MaleBirthRatio (MBR) calculated based on Bollywood movie dialogues**

| Time Period | Old | Mid | New |
|-------------|-----|-----|-----|
| MBR         | 73.9| 76.4| 54.5|

Old, mid, and new denote the time periods 1950–69, 1970–99, and 2000–20, respectively.
a specific genre, we consider 150 movies released after 1990. We confirm the genre of a movie through the genre lists or tags given by IMDB and Google. Figure 4 indicates that (1) the gender bias toward men for action movies is indeed a lot more pronounced than that in romantic movies; and (2) across both industries and movie genres, Hollywood action films exhibit the most bias.

**Diachronic word embeddings**

The meaning of words and the context in which they are used change over time. The language spoken in a community is representative of the cultural norms and customs followed in that region. Inspecting nearest neighbors of a given word in historical word embeddings (embeddings trained on different temporal slices of a large, longitudinal corpus) can reveal key insights. However, comparison of word vectors from different time periods requires that the vectors are aligned to the same coordinate axes. Hamilton et al. provide a robust multilingual approach to align diachronic word embeddings using orthogonal procrustes (estimating an orthogonal matrix that maps one set of points to another). We focus on the portrayal of women and men using diachronic embeddings. Figures 5A and 5B present the historical evolution of the nearest neighbors of the words *man* and *woman* in word embeddings trained on different temporal slices of our Bollywood and Hollywood data sets. We observe that the valence scores of the nearest neighbors for both genders across both movie industries show a similar pattern. The scores are the lowest during the 1970–99 period. The valence scores for the newer movies are better than the scores for the older movies. The dip in the valence scores during the period of 1970–99 in India can be ascribed to a social and cultural crisis influenced by an unstable political climate (assassinations of two prime ministers), two major wars between India and Pakistan, and a large overlap with the pre-economic liberalization period.

**Cloze tests**

In this section, we leverage recent advancements in language models to examine occupational stereotypes. When presented with a sentence (or a sentence stem) with a missing word, a cloze test is essentially a fill-in-the-blank task. For instance, in the following cloze test, *In the [MASK], it is very sunny, summer is a likely completion for the missing word. Given a cloze test, BERT, a well-known language model, outputs a series of tokens ranked by probability. In fact, in the above cloze test, the top three tokens (ranked by probability) predicted by BERT are summer, winter, and spring. Recent lines of research have explored BERT’s masked word prediction to (1) extract a knowledge base, (2) mine political insights and aggregate opinions, and (3) estimate linguistic quality.

We first fine-tune BERT models on our movie sub-corpora and investigate occupational stereotypes using the following two cloze tests:

1. A woman should be a [MASK] by occupation (denoted by cloze1);
2. A man should be a [MASK] by occupation (denoted by cloze2).

Next, we quantify the outputs of the models using a well-known lexicon of emotional valence ratings of nearly 14,000 English words to quantify the change of cloze test completions.
over time. The valence score of these words is presented on a scale of 1–10 with 10 indicating highly positive and 1 indicating highly negative. For example, the emotional valence scores of happy and sad are 8.47 and 2.10, respectively. For a given data set and a cloze test pair, we compute the average valence score of the top 10 completions ranked by probability (listed in square brackets in Table 2). We present further experimental details in the cloze test and free form completions section.

Our cloze test results are summarized in Table 2. We observe that completion results for both genders across both movie industries improve over time. We note that comparing completion results across genders using our lexicon may introduce certain biases. For instance, the valence scores for women are 5.42 and 7.09, respectively. We thus restrict our attention to comparing completion results across genders using our lexicon may introduce certain biases. For instance, the valence scores for women are 5.42 and 7.09, respectively. We thus restrict our attention to comparing completion results for both genders across both movie industries improve over time. However, for Bollywood, we notice that the rate of increase for women is substantially more pronounced than that for men. This observation aligns with the continual fight for gender equality in India and major movements that have mobilized voices for women’s right to work, financial independence, and marital laws.

Son preference: RQ 1.2
As already discussed in the gender attitudes and biases section, son preference in India is a well-documented phenomenon. Skewed sex ratio, female feticide, and higher child mortality rate for girls have attracted policymakers’ attention, leading to legal prohibition of prenatal sex discernment.

A popular Bollywood plot point is the introduction of a child into the family. Approximately, every one in ten collected movies had a scene involving birth of a child. We were curious to analyze when a child is born in a Bollywood movie, is it a boy or a girl? Let $N_w$ denote the number of times a dialogue talking about the baby’s gender $w$ appears in a corpus. We define male birth ratio (MBR) as follows: $\text{MBR} = \frac{N_{\text{boy}}}{N_{\text{boy}} + N_{\text{girl}}} \times 100$. Table 4 suggests that the family dynamics portrayed in Bollywood movies have shown considerable shift, with the MBR being 73.9 in older movies, to almost achieving parity (54.5) in newer movies.

Associating beauty with fair skin: RQ 1.3
RQ 1.3: Is beauty associated with fair skin in the movie dialogues describing women?

Previous studies have found that on social media and dating sites, women are often judged by their appearance, whereas men are mostly judged by their behavior. Skin color biases have been reported in the context of fairness beauty products in India, Indian arranged marriages, and surprisingly, political outcomes in India.

We first present our cloze test results with the probe “A beautiful woman should have [MASK] skin.” in Table 5. We note that while BERT$_{base}$ model’s top prediction is soft, all fine-tuned BERT models on the film corpora predict fair as the top choice. Figure 6 visualizes the nearest neighbors of beautiful in our aligned embedding spaces of Hollywood and Bollywood sub-corpora. As shown in Figure 6 and Table 5, the age-old affinity toward lighter skin in Indian culture is reflected through the consistent presence of fair among the nearest neighbors of all three Bollywood sub-corpora. Although our cloze tests indicate Hollywood also exhibits bias toward lighter skin color, our diachronic word embedding analysis reveals that possibly the bias is less pronounced than that in Bollywood.

Perception of dowry: RQ 1.4
As described in the gender attitudes and biases section, the dowry system involves a transaction of financial assets between money, debt, and loans. Jewellery, fees, and loans are considered dowry in different contexts.

| Religion | 1951 (%) | 1961 (%) | 1971 (%) | 1981 (%) | 1991 (%) | 2001 (%) | 2011 (%) |
|----------|----------|----------|----------|----------|----------|----------|----------|
| Hinduism | 84.1     | 83.4     | 82.7     | 82.6     | 81.5     | 80.5     | 79.8     |
| Islam    | 9.8      | 10.7     | 11.2     | 11.4     | 12.6     | 13.4     | 14.2     |
| Christianity | 2.3  | 2.4      | 2.6      | 2.4      | 2.3      | 2.3      | 2.3      |
| Sikhism  | 1.9      | 1.8      | 1.9      | 2.0      | 1.9      | 1.9      | 1.7      |
| Buddhism | 0.7      | 0.7      | 0.7      | 0.7      | 0.8      | 0.8      | 0.7      |
| Jainism  | 0.5      | 0.5      | 0.5      | 0.5      | 0.4      | 0.4      | 0.4      |
the bride’s family and the bridegroom’s family, with the latter being the recipient of the financial assets. Understandably, this retrograde system can influence son preference, as a girl child might be looked upon as financial burden. Despite legal prohibition since 1961, this practice has continued in India with several studies linking it to other social crises such as female feticide, domestic abuse and violence, and dowry deaths.

As shown in Figure 7, we observe that while nouns such as money, debt, jewellery, fees, and loan are the nearest neighbors in older films, indicating compliance to this practice, modern films exhibit non-compliance (e.g., guts and refused) and indicate some of the consequences of such non-compliance (e.g., divorce and trouble) in the form of nearest neighbors. Our findings align with a recent study based on a survey conducted among 4,603 women in Bihar (an Indian state in which the dowry has strong roots in tradition) that has reported positive changes in the society where the general attitude toward the dowry system has become negative.

Religion
RQ 2: How are religions perceived in movies? Can we gain an insight into the religious representation of a country through a film corpus spanning 70 years?
Perception of religion
Table 6 summarizes the religious distribution of six major religions in India according to the decennial censuses conducted since 1951. As indicated in Table 6, Hinduism and Islam are the two major religions in India, accounting for more than 90% of the country’s population. Two major partitions in the last 70 years faced by the Indian subcontinent have resulted in considerable religious turmoil and riots between these two communities.

The Central Board of Film Certification in India is a governing body that, along with giving each movie a certification, has the ability to remove offensive or controversial content, or in some extreme cases, it can completely ban films from being screened in theaters. With religion being a contentious topic in India, offensive terms surrounding it are also discouraged in films, and this has been constant throughout the years. To validate this hypothesis, we first look at the nearest neighbors of the word religion in the historical embeddings. Figure 8 indicates that religion is always accompanied with neutral or mild terms, and movie dialogues in Bollywood have stayed away from using extreme or hateful terms surrounding religion.

We next focus on the nearest neighbors of Hindu and Muslim in the historical embeddings and contrast our findings with prior research on aggregating social media perception of these two communities as presented in Palakodety et al. As shown in Figures 9A and 9B, we find that while negative words like ruthless, shameless, and traitor creeping up in newer movies might indicate religious polarization, words like terrorists found in social media data from Palakodety et al. are yet to surface among the nearest neighbors. Along with word embedding analysis, we analyze the BERT cloze tests for the probes (1) Hindus are [MASK] and (2) Muslims are [MASK]. For both probes, we do not notice completions such as terrorists or fools previously reported in Palakodety et al. This suggests...
that although recent social media analyses might indicate religious polarization, the film certification board has largely ensured movie content does not reflect such an extreme divide.

**Religious representation**

While the religious composition in India largely remained stable over the years, Table 6 indicates a slow decline in the population share of Hinduism and an increase of population share of Islam. In fact, the shifting nature of the demographic balance between the two major religions in India and its possible interpretations have a long history of use in political debates. We conduct a comprehensive analysis of the evolving nature of religious representation in Bollywood content through surname usage in our data set (Table 7 lists a set of highly frequent surnames occurring in Bollywood movies; details are presented in experimental procedures). Figure 10 contrasts the religion distribution obtained in movies with ground truth census data.

We note the following: (1) the distribution is more or less consistent with the census numbers; (2) representation for other religions has increased in recent years; and (3) the representation of Muslims is slightly less than the community’s population share.

Table 6 indicates the surnames of the doctors occurring in Bollywood movies. To retrieve these surnames from the subtitles, we employ a template-based approach, searching for keywords like "Dr." and "doctor" from our corpus. While a broader religious representation is observed in our overall results, we find that the representation for the medical profession is quite skewed, with a large number of surnames being Brahmins (the uppermost caste in the Hindu caste system in India).

**Attitudes and biases toward geographic regions**

**RQ 3:** How has geographic representation evolved over time in Bollywood content? Which geographic areas have been consistently under-represented in the Mumbai film industry?

As already discussed (in the attitudes and biases toward geographic regions section), similar to religion, regional politics has been a recurrent theme in Indian political discourse. From the beginning, Bollywood has had its roots in Mumbai, and Delhi is India’s capital. Hence, it is not surprising that these cities are mentioned heavily across all time periods (see Table 9). Figures 11A and 11B compare the geographic representations in the most recent 20 years with the rest of our corpus spanning 1950–99. We observe that initially based out of major hotspots of Delhi, Goa, and cities like Mumbai, recent Bollywood content...
is geographically more diverse and inclusive. However, a key point we highlight in Figure 11C is that, in line with prior research on under-representation of northeastern states in news content, there is severe under-representation of these northeastern states in Bollywood content. In fact, there have been zero mentions of the states of Arunachal Pradesh, Meghalaya, and Mizoram in over 700 movies across 70 years.

**Economic signals and national priorities**

**RQ 4: Can we extract economic signals through popular film dialogues? Can we track evolving national priorities from popular entertainment?**

**Economic signals**

By looking at a popular entertainment corpus of a developing nation, we are able to showcase the evolution of gender bias, evolving attitudes toward social evils, and geographic and religious representations. Can we detect economic signals as well from the Bollywood dialogues? 100 rupees in 1958 is equivalent to 8,117.22 rupees in 2020 (https://www.inflationtool.com/indian-rupee). We seek to understand whether language models can capture these noisy signals. GPT–2, a popular language model with more than 100 million parameters, has achieved state-of-the-art results for text completion, zero shot transfer learning, etc. GPT–2 has been widely used for generating free form text to create artificial newsletters, poems, etc. (https://www.gwern.net/GPT-2). We noticed that the most common dialogues expressing monetary figures or large amounts of money were generally associated with ransom. For example, a sample dialogue is “We have kidnapped your kid; the ransom amount is 2 million rupees.”

To analyze the historical trends, we fine-tune GPT–2 on three Bollywood sub-corpora, each belonging to films from different time periods, for the end goal of free form text completion. On these fine-tuned models, we input the sentence “The ransom amount is [MASK]” and analyze the generated text by the model. Table 10 showcases the average amount across 100 generated samples from the fine-tuned models. We note that while our predicted values overestimate the inflation rate, the ransom amounts capture the general increasing pattern and have increased significantly over time.

**Evolving national priorities**

Similar to BERT’s cloze test applications to uncover occupational stereotypes and bias toward fair skin color, following Palakodety et al. and KhudaBukhsh et al., we employ BERT to analyze evolving national priorities using the following two cloze tests:

1. The biggest problem in India is [MASK] (for Bollywood).
2. The biggest problem in America is [MASK] (for Hollywood).

We observe that the dynamic political conditions are reflected in the completion results in Tables 11 and 12 (e.g., Kashmir, Pakistan, and Russia). We also note that the list of ongoing problems in the United States contains the major issue in the 2020 election: racism.

**DISCUSSION**

In this paper, we analyzed how social biases and subtle gender biases get reflected on diachronic corpora of popular entertainment. Our research indicates that our NLP methods are capable of uncovering important social signals. Some of the findings in our papers are anecdotally known or expected. For instance, it is not difficult to envision that award-winning, critically acclaimed world movies may exhibit more progressive attitudes toward gender equality than Bollywood potboilers. And, it is not surprising that male-dominated action movies would exhibit larger gender bias than romantic films. Many of our research questions have been explored in prior social science literature. However, such efforts are typically limited to a handful of films (see, e.g., Dimitrova, Rao, and Lundberg). However, what we offer here is a comprehensive, quantitative, and large-scale analysis of these research questions using a suite of cutting-edge NLP techniques in a synergistic way. Our methods thus allow us to measure more objectively and quantitatively. They also allow precise tracking of change in these biases over time.

Large-scale analyses of texts are intractable without automated methods. We present statistical, automated analysis of movies at scale and across time that gives us a finer probe for understanding the cultural themes implicit in these films. In our

| Old | Mid | New |
|----|----|----|
| Bombay/Mumbai (51) | Bombay/Mumbai (68) | Bombay/Mumbai (83) |
| Delhi (27) | Delhi (45) | Delhi (52) |
| Kolkata/Calcutta (23) | Kolkata/Calcutta (18) | Amritsar (9) |
| Lucknow (14) | Lucknow (12) | Bangalore/Bengaluru (9) |
| Madras/Chennai (10) | Simla/Shimla (12) | Kolkata/Calcutta (9) |
| Agra (6) | Madras/Chennai (10) | Pune (8) |
| Srinagar (6) | Pune (8) | Lucknow (7) |
| Simla/Shimla (6) | Bangalore/Bengaluru (7) | Hyderabad (6) |
| Mathura (5) | Nagpur (6) | Madras/Chennai (6) |

Table 9. City mentions in movies from our Bollywood corpus. Old, mid, and new denote the time periods 1950–1969, 1970–1999, and 2000–2020, respectively.
study, we restrict ourselves to 100 movies per decade. However, the methods presented in this paper can easily scale up to 1,000 or more movies as long as we have English subtitles for them. Also, the same NLP tools might be used to rapidly analyze hundreds or thousands of books, magazine articles, radio transcripts, or social media posts.

Not all the uncovered biases are known. For instance, our analyses indicate that babies born inside popular Bollywood films exhibit a skewed gender distribution. This gender disparity has evolved (and improved) over time. Without a large-scale analysis, these types of insights are hard to obtain.

When contrasted with real-world data, we observe that our findings in movies somewhat align to real-world data. However, sometimes movies show better numbers than present in the society, and sometimes the numbers are worse. For example, in our analyses, we observe that the representation of Muslims in Bollywood movies is slightly less than the community’s population share. In the case of son preference, our results indicate an improving gender ratio (MBR) over time. However, in reality, the 2011 census recorded an all-time-low CSR. Similarly, our economic signal overestimates the actual inflation rate while retaining the general increasing pattern. Our geographic representation results also corroborate the historical under-representation of northeastern states in other media.

With the advent of the multiplex era in Indian cinema, different movies can offer different prices. Hence, if a progressive idea has enough takers, movies exhibiting that progressive idea can generate substantial revenue. When certain numbers are better in movies than in the society, we hypothesize that this is a signal that society could be receptive to the idea, and that we might see real numbers in society improve over the coming years. (In fact, a recent study has revealed that some of the most imbalanced districts in India are on their way to recording better CSR numbers.) In contrast, when numbers in movies are worse than society, this could be a signal that the society is less receptive to the progressive idea.

Our results demonstrate that societal changes do get reflected in popular content. But does popular entertainment also influence the society in turn? A recent movie on acid attack, Chhapak, was inspired from a true story of an acid attack survivor who set up an NGO and was a recipient of the International Women of Courage award. Her biopic and her initiative of Stop Acid Sale when released, triggered regulatory legislation that made it difficult to buy certain types of acids without legal authorization.
Devising NLP methods to identify how popular entertainment influences society will be a worthy future research challenge.

Limitations
Our study has several limitations. For some of our experiments, we build upon pre-trained language models trained on vast amounts of texts. While our experiments on diachronic word embeddings and WEAT train the word embeddings from scratch, our experiments using BERT and GPT-2 are trained on top of pre-trained models. Recent works have indicated that these models have a wide range of biases that reflect the texts on which they were originally trained, and which may percolate to downstream tasks.86

Our study focuses on linguistic signals obtained from English subtitles of popular films. However, films have a strong visual component, and linguistic signals may not be able to capture biases present in the visual medium. For example, Dhoom 2, a commercially successful movie released in 2006, belongs to our data set. In one of the scenes in this film, Ali, one of the main characters in the film, fantasizes a future family with Monali, another character in the film. In this sequence with screen time of less than 10 s and without any dialogue, Ali’s dream family indicates that the two of them (Ali and Monali) have four children, where all of them are sons. Our analysis of son preference relying solely on movie dialogues will be unable to capture this subtle visual signal hinting deep-seated son preference. Our study will substantially benefit from a multi-modal analysis considering both linguistic and visual signals.

Finally, Bollywood represents a fraction of Indian cinema. There are many regional language movie industries (e.g., Kannada, Telugu, Tamil, Malayalam, Bengali, and Marathi) that heavily contribute to Indian cinema. According to the Central Board of Film Certification report (http://www.filmfed.org/downloads/Language-wise-Region-2018-19-26062019.pdf), out of the 1,966 movies certified by the board in 2019, only 495 movies (24.92%) came from the Mumbai film industry. Hence, extending our study to regional movies will strengthen our analyses further. Also, the location of Mumbai and linguistic barrier may influence some of our analyses. For example, in our geographic representation analysis, we found minuscule representation of several northeastern states. While these states have documented under-representation in other media such as the print news medium, it is possible that linguistic barrier and geographic distance were potential factors.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact
Requests for data and requests for additional information should be directed to the lead contact, Ashiqur R. KhudaBukhsh (axk@seas.rut.edu).

Materials availability
This study did not generate physical materials.

Data and code availability
Data and code are available at https://github.com/kunaikhadilkar/CellPatternsBollywood.

Data set

We construct the following two data sets of movie subtitles.

1. Bollywood movies, \(D_{\text{bolly}}\): We consider 100 top-grossing movies for each decade spanning 1950–2020 (7 decades, and overall, 700 movies). We retrieve English subtitles86 for each of these 700 movies.

2. Hollywood movies, \(D_{\text{holly}}\): Similar to Bollywood movies, we consider 100 top-grossing movies from each of the seven decades (700 total films). Overall, \(D_{\text{bolly}}\) and \(D_{\text{holly}}\) consist of 1.1M dialogues (6.2M tokens) and 1M dialogues (5.4M tokens), respectively.

In several experiments, we divide our corpus into three temporal buckets, presented in Tables 13 and 14. Our choice of separation points in the timeline is guided by the global emergence of counter-culture in the late 60s and early 70s87 and the rapid rise of multiplex culture in Indian cinema.88 We understand that several other reasonable binning choices exist. In the supplemental information (Table S3), we have shown experimental results that indicate that our qualitative claims remain unchanged with an alternative binning.

We further collect 150 movies that have been nominated for the Best International Feature Film award 1970 onward at the Oscars (https://www.oscars.org) for a subset of our analyses.

Cloze test and free form completions

In our experiments with language models, we use both BERT and GPT-2 because of their different capabilities. For instance, BERT is particularly well-suited for cloze tests, and there exists substantial literature where BERT has

| Table 11. Cloze test results for The biggest problem in India is [MASK] |
|-----------------------------|-----------------------------|-----------------------------|
| **BERT\textsubscript{base}** | **BERT\textsubscript{pre}** | **BERT\textsubscript{post}** |
| corruption (0.034), poverty (0.020), malaria (0.019), pollution (0.012), hunger (0.012), terrorism (0.009), unemployment (0.008), drought (0.008), famine (0.007), war (0.003), tourism (0.001) | poverty (0.078), love (0.072), war (0.067), hunger (0.049), unemployment (0.043), India (0.042), famine (0.029), money (0.023), marriage (0.012), education (0.011), Kashmir (0.009) | poverty (0.074), Pakistan (0.072), Kashmir (0.053), terrorism (0.051), corruption (0.037), India (0.031), drugs (0.021), dowry (0.016), unemployment (0.014), hunger (0.009), rape (0.006) |

| Table 12. Cloze test results for The biggest problem in America is [MASK] |
|-----------------------------|-----------------------------|-----------------------------|
| **BERT\textsubscript{base}** | **BERT\textsubscript{pre}** | **BERT\textsubscript{post}** |
| poverty (0.076), corruption (0.072), unemployment (0.061), crime (0.045), terrorism (0.042), racism (0.027), pollution (0.021), hunger (0.016), war (0.012), cancer (0.009), inequality (0.003) | war (0.092), poverty (0.083), money (0.062), unemployment (0.053), slavery (0.051), immigration (0.045), alcoholism (0.041), education (0.032), imperialism (0.023), Russia (0.023), hunger (0.019) | poverty (0.088), slavery (0.082), immigration (0.078), unemployment (0.073), money (0.071), war (0.065), racism (0.053), hunger (0.024), communism (0.016), America (0.011), education (0.006) |
been deployed for this task. On the other hand, for our economic signal mining task, the ransom money can be specified as a free form text (e.g., 0.5 million dollars or 500 grands). GPT-2 is particularly well-suited for free form text completions. Moreover, the outputs of GPT-2 are non-deterministic, which allows us to conduct multiple runs of the same experiment and compute confidence intervals.

We follow the standard preprocessing steps recommended to fine-tune BERT language model. For our task, we use the bert-base-uncased pre-trained English model, with the following parameter details: 12 transformer layers, hidden state length of 768, 12 attention heads, and 110M overall parameters (denoted as BERTbase). The pre-trained model is fine-tuned on the target corpus using the training parameters showcased below.

- batch size: 16
- maximum sequence length: 128
- maximum predictions per sequence: 20
- fine-tuning steps: 10,000
- warmup steps: 10
- learning rate: 2e–5

For fine-tuning the language model for free form text completion tasks, we use the smallest GPT-2 model with 124M parameters, trained for 10,000 steps.

**WEAT**

We follow the experimental protocols identical to those specified in Van Miltenburg, the paper that introduced this technique. Following Van Miltenburg, we train the sub-corpora using GloVe embeddings. We consider two equal-sized sets of occupations, S1 and S2, and two sets of attribute words, A1 and A2.

The similarity of two words, say x and y, is given by calculating the cosine similarity of the corresponding word embeddings, \( \cos(w_x, w_y) \). As given in Van Miltenburg, the differential association of a word \( c \) with word sets \( A_1 \) and \( A_2 \) is given by the following:

\[
g(c, A_1, A_2, w) = \max_a A_1 \cos(w_x, w_y) - \max_b A_2 \cos(w_x, w_y)
\] (Equation 1)

Next, the WEAT score is calculated as:

\[
B_{\text{weat}}(w) = \frac{\max_a S_1 g(s, A_1, A_2, w) - \max_b S_2 g(s, A_1, A_2, w)}{S_1 g(s, A_1, A_2, w)}
\] (Equation 2)

The occupation sets, \( S_1 \) and \( S_2 \), are taken from Bolukbasi et al. A1 = [maestro, skipper, protege, philosopher, captain, architect, financier, warrior, broadcaster, magician, pilot, boss].

\( S_2 \) = [homemaker, nurse, receptionist, librarian, socialite, hairdresser, nanny, bookkeeper, stylist, housekeeper, designer, counselor].

The attribute word sets \( A_1 \) and \( A_2 \) are \( A_1 = \{he, man, male\} \) and \( A_2 = \{she, woman, female\} \).

In order to investigate if the WEAT results are mere artifacts of the embedding model or not, we conduct the following experiment. We modify the corpora by randomly flipping the gendered pronouns in the corpus and then train the embedding models and recompute the WEAT scores. We find that the WEAT score is close to zero (0.0158 ± 0.0039). Thus the bias present in the model is minimal, and our obtained results are reliable.

**Aligning diachronic word embeddings**

We use a robust multilingual approach to align diachronic word embeddings using orthogonal Procrustes (estimating an orthogonal matrix that maps one set of points to another) as described in Hamilton et al. We follow the same method to align different sub-corpora for Bollywood and Hollywood. For each sub-corpora listed in Tables 13 and 14, we train word2vec with SGNS (skip-gram with negative sampling) to obtain word embeddings. Let \( \mathbf{V}_t \) be the matrix of word embeddings learnt for period \( t \) for vocabulary \( \mathcal{V} \). Following Mikolov, we align the word embeddings using the top 10,000 common tokens present across time periods \( t \) and \( t+1 \) by optimizing:

\[
R' = \arg\min_{R \in \mathbb{C}^{d \times d}} \| \mathbf{QW}' - W'(t+1) \|_F
\]

where \( R' \in \mathbb{C}^{d \times d} \). Note that, our choices of word embeddings for our experiments to compute WEAT score and to align diachronic word embeddings differ because in all our experiments, for each of the techniques, we have followed experimental protocols identical to those specified in the papers that introduced these techniques.

**Son preference**

We retrieve the dialogues talking about childbirth using a template-based approach, by searching for the following keywords and phrases: birth, baby, pregnant, pregnancy, congratulations, “It’s a boy,” and “It’s a girl.” We annotated the retrieved dialogues related to childbirth and performed a temporal analysis.

**Religious representation**

437 surnames appearing in the movies from our corpus (e.g., Mrs. Kapoor, Mr. Khan, etc.) are annotated manually by two annotators, with each surname given one label from the list of labels: Hindu, Muslim, Sikh, Christian, Parsi, or multiple. Since almost all Jain surnames are also highly prominent Hindu surnames (e.g., Mehta, Chopra), and unlike several major religions, Buddhism does not have easy-to-distinguish last names, we exclude Jainism and Buddhism in our analysis. The annotators achieved a Cohen’s \( \kappa \) score of 0.8879, indicating high inter-rater agreement. The discrepancies were resolved by consulting literature (e.g., tracking biographies of prominent personalities having a specific last name). A random sample of prominent Indian surnames annotated with religions is presented in the supplemental information (Table S2).

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at [https://doi.org/10.1016/j.patter.2021.100409](https://doi.org/10.1016/j.patter.2021.100409).

**ACKNOWLEDGMENTS**

We are thankful to Parichay Patra for his insightful feedback. We express our gratitude to the anonymous reviewers for their thoughtful, valuable, and constructive comments.

**AUTHOR CONTRIBUTIONS**

A.K. conceived the idea, approach, and research questions; K.K. collected the data set and performed experiments; A.K. and T.M. designed the experiments; A.K., K.K., and T.M. wrote the manuscript.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.
REFERENCES

1. Khadilkar, K., and KhudaBukhsh, A.R. (2021). An unfair affinity toward fairness: characterizing 70 years of social biases in Bollywood (student abstract). In Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021 [AAAI Press], pp. 15815–15814. https://ojs.aaai.org/index.php/AAAI/article/view/17903.

2. Jaffrelot, C. (2010). Religion, Caste, and Politics in India (Primus Books).

3. Zargar, H. (2020). How Bollywood furthers India’s Hindu nationalism. https://www.newframe.com/how-bollywood-furthers-indias-nationalism/.

4. Dimitrova, D. (2010). Religion and gender in bollywood film. In Religion in Literature and Film in South Asia (Springer), pp. 69–81.

5. Ramaswamy, S. (1997). Passions of the Tongue: Language Devotion in Tamil India, 1891–1970, Vol. 29 (Univ of California Press).

6. Rao, L. (1989). Woman in Indian films: a paradigm of continuity and change. Media, Cult. Soc. 11, 443–458.

7. Tvenge, J.M., Campbell, W.K., and Gentile, B. (2012). Male and female pronoun use in US books reflects women’s status, 1900–2008. Sex Roles 67, 488–493.

8. Lahoti, R., and Swaminathan, H. (2016). Economic development and women’s labor force participation in India. Feminist Econ. 22, 168–195.

9. Garg, N., Schiebinger, L., Jurafsky, D., and Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. Proc. Natl. Acad. Sci. U S A 115, E3635–E3644.

10. R. Pande, A. Mahotra, S. Mathur, M. Mehta, A. Malhotra, M.A. Lycette, S.D. Kambou, V. Magar, J. Gay, H. Lary, et al., Son Preference and Daughter Neglect in India. International Center for Research on Women.

11. Bhat, P.N.M., and Zavier, A.J.F. (2003). Fertility decline and gender bias in Northern India. Demography 40, 637–657. http://www.jstor.org/stable/1515201.

12. R. Echa´ varri, Gender Bias in Sex Ratio at Birth: The Case of India.Universidad Pública de Navarra

13. Mukherjee, S.S. (2013). Women’s empowerment and gender bias in the politics of olympic sport in postcolonial India. Int. J. Hist. Sport 23, 1217–1231.

14. N. Shevde, All’s fair in love and cream: a cultural case study of fair & lovely skin-care products.

15. Nagar, I. (2018). Why dowry continues to prevail. J. Develop. Areas, 71–93.

16. Ahmad, N. (2008). Dowry deaths (bride burning) in India and abetment of suicide: a socio-legal appraisal. JE Asia Int’l L. 1, 275.

17. Srinivasan, P., and Lee, G.R. (2004). The dowry system in Northern India: women’s attitudes and social change. J. Marriage Fam., 66, 1108–1117.

18. Pethe, V.P. (1973). Hindus, muslims and the demographic balance in India. Econ. Polit. Weekly 8, 75–78.

19. Talbot, I., and Singh, G. (2009). The Partition of India (Cambridge University Press).

20. Iyer, S., and Shrivastava, A. (2018). Religious riots and electoral politics in India. J. Develop. Econ. 131, 104–122.

21. Majumdar, B. (2006). When north–south fight, the nation is out of sight: the politics of olympic sport in postcolonial India. Int. J. Hist. Sport 23, 1217–1231.

22. Chongkhi, H. (2017). Portrayal of northeast India in mainstream media: a case of underrepresentation and misinterpretation. Int. J. Res. Social Sci. 7, 613–623.

23. Palakodety, S., KhudaBukhsh, A.R., and Carbonell, J.G. (2020), Mining insights from large-scale corpora using fine-tuned language models. In Proceedings of the Twenty-Fourth European Conference on Artificial Intelligence (ECAI-20), pp. 1890–1897.

24. R. KhudaBukhsh, R. Sarkar, M.S. Kamlet, T.M. Mitchell, Fringe news networks: dynamics of US news viewership following the 2020 presidential election, arXiv:2101.10112.

25. Van Miltenburg, E. (2016). Stereotyping and bias in the flickr30k dataset. arXiv, 1605.06083.

26. Caliskan, R., Bryson, J.J., and Narayanan, A. (2017), Semantics derived automatically from language corpora contain human-like biases. Science 356, 183–186.

27. Lewis, M., and Lupyany, G. (2020). Gender stereotypes are reflected in the distributional structure of 25 languages. Nat. Hum. Behav. 4, 1–8.

28. Prates, M.O., Avelar, P.H., and Lamb, L.C. (2019). Assessing gender bias in machine translation: a case study with google translate. Neural Comput. Appl. 32, 1–19.

29. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al., (2020), language models are few-shot learners. In Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6–12, virtual https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac1426f6a-Abstract.html.

30. Lucy, L., Demszy, D., Bromley, P., and Jurafsky, D. (2020). Content analysis of textbooks via natural language processing: findings on gender, race, and ethnicity in Texas U.S. history textbooks. AERA Open 6. https://doi.org/10.1177/2332858420940312.

31. Gala, D., Khursheed, M.O., Lerner, H., O’Connor, B., and Iyyer, M. (2020). Analyzing gender bias within narrative tropes. arXiv, 2011.00092.

32. Bolukbasi, T., Chang, K.-W., Zou, J.Y., Saligrama, V., and Kalai, A.T. (2016). Man is to computer programmer as woman is to homemaker? debunking word embeddings. Adv. Neural Inf. Process. Syst. 29, 4349–4357.

33. Van Miltenburg, E. (2016). Stereotyping and bias in the flickr30k dataset. arXiv, 1605.06083.

34. Caliskan, R., Bryson, J.J., and Narayanan, A. (2017), Semantics derived automatically from language corpora contain human-like biases. Science 356, 183–186.

35. Lewis, M., and Lupyany, G. (2020). Gender stereotypes are reflected in the distributional structure of 25 languages. Nat. Hum. Behav. 4, 1–8.

36. Prates, M.O., Avelar, P.H., and Lamb, L.C. (2019). Assessing gender bias in machine translation: a case study with google translate. Neural Comput. Appl. 32, 1–19.

37. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al., (2020), language models are few-shot learners. In Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6–12, virtual https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac1426f6a-Abstract.html.

38. Lucy, L., Demszy, D., Bromley, P., and Jurafsky, D. (2020). Content analysis of textbooks via natural language processing: findings on gender, race, and ethnicity in Texas U.S. history textbooks. AERA Open 6. https://doi.org/10.1177/2332858420940312.

39. Gala, D., Khursheed, M.O., Lerner, H., O’Connor, B., and Iyyer, M. (2020). Analyzing gender bias within narrative tropes. arXiv, 2011.00092.
45. Khan, S., and Taylor, L. (2018). Gender policing in mainstream Hindi cinema: a decade of central female characters in top-grossing bollywood movies. Int. J. Commun. 12, 22.

46. Madaan, N., Mehta, S., Agrawala, T.S., Mahotra, V., Aggarwal, A., Gupta, Y., and Saxena, M. (2018). Analyze, detect and remove gender stereotyping from bollywood movies. In Conference on Fairness, Accountability and Transparency, FAT 2018, 23–24 February 2018, New York, NY, USA, Vol. 81 of Proceedings of Machine Learning Research, S.A. Friedler and C. Wilson, eds. (PMLR), pp. 92–105. http://proceedings.mir.press/v81/madaan18a.html.

47. S. Sheth, G. Jones, M. Spencer, Bollywood, Skin Color and Sexism: The Role of the Film Industry in Emboldening and Contesting Stereotypes in India after Independence. Harvard Business School

48. Mishra, N. (2015). India and colorism: the finer nuances. Wash. U. Glob. Stud. L. Rev. 14, 725.

49. Breakthrough (2019). 8 examples of nerve-wracking misogynist dialogues in bollywood. https://nbreakthrough.org/misogynist-dialogues-bollywood/.

50. Thesecondangle (2020). Top 10 sexist dialogues in bollywood. https://thesecondangle.com/sexist-dialogues-bollywood/.

51. Popxo (2019). Stupid and sexist: 20 WTF dialogues that prove bollywood is unfair towards women. https://www.popxo.com/2019/07/20-bollywood-dialogues-that-are-worst-examples-of-sexism/.

52. Wanner, B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. Behav. Res. Methods 45, 1191–1207.

53. Sendén, M.G., Sikström, S., and Lindholm, T. (2015). “She” and “He” in news media messages: pronoun use reflects gender biases in semantic contexts. Sex Roles 72, 40–49.

54. Taylor, W.L. (1953). “Cloze procedure”: a new tool for measuring readability. Journalism Q. 30, 415–433.

55. Smith, N., and Levy, R. (2011). Cloze but no cigar: the complex relationship between cloze, corpus, and subjective probabilities in language processing. Proceedings of the Annual Meeting of the Cognitive Science Society, 33.

56. Petroni, F., Rocktaschel, T., Riedel, S., Lewis, P., Bakhtin, A., Wu, Y., and Miller, A. (2019). Language models as knowledge bases? In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (Association for Computational Linguistics), pp. 2463–2473. https://doi.org/10.18653/v1/D19-1250.

57. Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (NAACL-HLT), pp. 4171–4186.

58. Hamilton, W.L., Leskovec, J., and Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (Association for Computational Linguistics), pp. 1489–1501. https://doi.org/10.18653/v1/P16-1141. https://www.aclweb.org/anthology/P16-1141.

59. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. Language models are unsupervised multitask learners. https://d4mucfpksyw.cloudfront.net/better-language-models/language-models.pdf.

60. Bhardwaj, R. (2007). On the persistent political under-representation of Muslims in India. L. Ethics Hum. Rights 1, 76–133.

61. Greenwald, G., and Banaji, M.R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychol. Rev. 102, 4.

62. Xie, J.Y., Ferreira Pinto Junior, R., Hirst, G., and Xu, Y. (2019). Text-based inference of moral sentiment change. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (Association for Computational Linguistics), pp. 4654–4663. https://doi.org/10.18653/v1/D19-1472. https://www.aclweb.org/anthology/D19-1472.

63. Hardgrave, R.L., Jr. (1985). India in 1984: confrontation, assassination, and succession. Asian Surv. 25, 131–144.

64. Kaarthikenyan, D.R., and Raju, R. (2008). Rajiv Gandhi Assassination (Sterling Publishers Pvt. Ltd.).

65. Schofield, V. (2010). Kashmir in Conflict: India, Pakistan and the Unending War (Bloomsbury Publishing).

66. Bose, S. (2009). Kashmir: Roots of Conflict, Paths to Peace (Harvard University Press).

67. Pedersen, J.D. (2000). Explaining economic liberalization in India: state and society perspectives. World Develop. 28, 265–282.

68. N. Kassner, P. Dufter, H. Schütze, Multilingual LAMA: investigating knowledge in multilingual pretrained language models, arXiv:2102.00894.

69. Sarkar, R., Mahinder, S., and KhudaBuksh, A.R. (2020). The non-native speaker aspect: Indian English in social media. In Proceedings of the Sixth Workshop on Noisy User-Generated Text, W-NUT@EMNLP 2020 Online, November 19, 2020 (Association for Computational Linguistics), pp. 61–70. https://doi.org/10.18653/v1/2020.wnut-1.9.

70. Nayak, P., and Mahanta, B. (2012). Women empowerment in India. Bull. Polit. Econ. 5, 155–183.

71. Chen, M. (1995). A Matter of Survival: Women’s Right to Employment in India and Bangladesh, Vol. 38 (Clarendon Press).

72. Goyal, M., and Parkash, J. (2011). Women entrepreneurship in India-problems and prospects. Int. J. Multidisc. Res. 1, 195–207.

73. Nigam, S. (2005). Understanding Justice Delivery System from the Perspective of Women Litigants as Victims of Domestic Violence in India (Specifically in the Context of Section 498-A, IPC) (Nigam Shalu), Occasional Paper (35).

74. Fiore, A.T., Taylor, L.S., Mendelsohn, G.A., and Hearst, M. (2008). Assessing attractiveness in online dating profiles. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 797–806.

75. Rose, J., Mackey-Kallis, S., Styles, L., Barry, K., Biagini, D., Hart, C., and Jack, L. (2012). Face it: the impact of gender on social media images. Commun. Q. 60, 588–607.

76. Otterbacher, J. (2015). Linguistic bias in collaboratively produced biographies: crowdsourcing social stereotypes? In Ninth International AAAI Conference on Web and Social Media.

77. Dlova, N.C., Hamed, S.H., Tsoka-Gwegweni, J., and Grobler, A. (2015). Skin lightening practices: an epidemiological study of South African women of African and Indian ancestries. Br. J. Dermatol. 173, 2–9.

78. Chattopadhyay, S. (2019). Fair-Unfair: Prevalence of Colorism in Indian Matrimonial Ads and Married Women’s Perceptions of Skin-Tone Bias in India, Ph.D. thesis (Bowling Green State University).

79. Madhukalya, A. (2020). These dialogues from bollywood blockbusters are so sexist that you’ll want to pull your hair out. https://www.huffingtonpost.in/2016/09/26/these-dialogues-from-bollywood-blockbusters-are-so-sexist-that-youll-want-to-pull-your-hair-out/.

80. Olcott, M. (1944). The caste system of India. Am. Sociol. Rev.

81. LaFeber, W., and Abbott, B. (1972). America, Russia, and the Cold War, 1945-1975 (Wiley).

82. M. Lundberg, Bollywood images: the illusions and realities of arranged marriages, weddings, dowries, and attitudes toward the girl child in the lives of women in India, Augsburg Honors Rev. 6.

83. Diamond-Smith, N., Saikia, N., Bishai, D., and Canudas-Romo, V. (2020). What has contributed to improvements in the child sex ratio in select districts of India? a decomposition of the sex ratio at birth and child mortality. J. Biosoc. Sci.

84. Firstpost (2014). Acid attack victim laxmi to receive international women of courage award - India News. https://www.firstpost.com/india/acid-attack-victim-laxmi-to-receive-international-women-of-courage-award-1418519.html.

85. Bender, E.M., Gebu, T., McMillan-Major, A., and Shmitchell, S. (2021). On the dangers of stochastic parrots: can language models be too big? In
86. P. Lison, J. Tiedemann, Opensubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles. Universitetsbiblioteket i Oslo
87. Greene, D. (2016). Rock, Counterculture and the Avant-Garde, 1966-1970: How the Beatles, Frank Zappa and the Velvet Underground Defined an Era (McFarland).
88. Athique, M., and Hill, D. (2007). Multiplex cinemas and urban redevelopment in India. Media Int. Aust. 124, 108–118.
89. Pennington, J., Socher, R., and Manning, C. (2014). GloVe: global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (Association for Computational Linguistics), pp. 1532–1543. https://doi.org/10.3115/v1/D14-1162. https://aclanthology.org/D14-1162.
90. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems, pp. 3111–3119.