Language of Outgroup Prejudice: Revealing Perceptions towards 60 U.S Immigrant Groups from Language Patterns

Ying Li^{1,2} and Thomas Hills²

¹ Max Planck Institute for Human Development, Berlin, Germany

² Department of Psychology, University of Warwick, UK

Author Note

Ying Li, Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany, and Department of Psychology, University of Warwick, UK;
Thomas Hills, Department of Psychology, University of Warwick, UK.

Word count: 6349

The authors thank Susannah Goss for her editing.

There is no competing interesting in this study.

Author contributions:

Y.Li developed the study concept. Both authors contributed to the study design. Data collection and analysis were performed by Y.Li. Both authors drafted the manuscript. Both authors approved the final version of the manuscript for submission.

Correspondence concerning this article should be addressed to Ying Li, Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: li@mpib-berlin.mpg.de

Code, visualizations of LDA topic model can be retrieved from <https://osf.io/tvkn3/>.
Data can be accessed from <https://catalog.ldc.upenn.edu/LDC2008T19>

Abstract

Migration is a central activity of human cultural history and brings diverse groups into contact with one another. In this study, we present cultural representations of 60 US immigrant groups by analyzing language patterns associated with them in a corpus of 1.8 million newspaper articles over a 20-year period. Based on construal level theory, we first operationalize perceived social distance towards outgroups as language concreteness and sentiment as language valence. We found that concrete language, as an indicator of close social contact, is strongly associated with positive sentiment. Next, we used a topic model to identify 15 topics (e.g., crime, politics, food, arts, and education) reflecting cultural representations of immigrant identity that may drive positive and negative sentiment. The present study, by analyzing language patterns of outgroups, provides an opportunity to evaluate prejudice and stereotype at large societal scale.

Keywords: *immigration, natural language processing, intergroup contact theory, social proximity*

Language of Outgroup Prejudice: Revealing Perceptions towards 60 Immigrant Groups from Language Patterns

In 2015, there were approximately 250 million immigrants worldwide (World Bank, 2017). Although immigrants bring innovation, skilled labor, investment, and cultural diversity to host countries (Borjas, 1990; Carens, 2013; Skeldon, 2014), they are commonly perceived as untrustworthy outsiders (Alexander, Brewer, & Herrmann, 1999; Cuddy, Fiske, & Glick, 2007; Cuddy et al., 2009; Peabody, 1985; Poppe, 2001). Negative public attitudes toward immigrants can have serious consequences for policy and decision making, with Brexit being a case in point. Understanding the mechanisms that form and shape public perceptions of outgroups is a crucial first step toward providing solutions that enhance outcomes for immigrants as well the populations into which they migrate.

Outgroup negativity is deeply rooted in the basic human propensity for social-categorical thinking (Allport, 1954; Brewer, 1979; Tajfel, 1982). Ultimate attribution error, for example, is the propensity to explain others' negative behaviors as resulting from dispositional properties of a categorically defined outgroup, but their positive behaviors as the result of idiosyncratic situational factors (Pettigrew, 1979). What is remarkable is that outgroup status and sentiment is decidedly flexible. Laboratory analogs of group formation—often called minimal group paradigms—have demonstrated that the minimum condition for intergroup bias is categorization into a group. The criteria for that categorization can be totally arbitrary, such as a preference for Kandinsky over Klee (Tajfel et al., 1971). In Sherif et al.'s (1961) Robbers' Cave experiment, boys at a camp were assigned to groups at random. Increasing levels of prejudice and hostility toward outgroup members were observed over a period of weeks. When the groups worked toward a common goal together, however, the boundaries between groups were rapidly broken down.

According to intergroup contact theory (Allport, 1954), fostering direct interaction between groups is an effective way to resolve outgroup prejudice because it reduces perceived social distance. A meta-analysis of more than 500 studies found that increased intergroup contact reduced prejudice in 94% of independent samples (Pettigrew & Tropp, 2006; also see a more recent meta-analysis: Paluck, Green & Green, 2018). Generally speaking, social distance fuels dispositional inference and prejudice (Jones & Nisbett, 1972). For example, intergroup contact plays a substantial role in explaining the rural–urban divide in perceptions about immigrants, whereby rural populations who have little contact with

immigrants tend to have more negative attitudes toward immigration than do urban populations who interact with them regularly (Fennelly & Federico, 2008).

Intergroup contact theory makes more nuanced demands of intergroup interactions than contact simply having occurred, however. The interaction must have prosocial qualities, such as equal status, cooperation, and social approval, all of which reduce perceived social distance (Allport, 1954). Replications of Bogardus's (1927) study of contact among the most common social groups in the United States have found that the mean level of social distance toward all social groups has decreased over time (Parrillo & Donoghue, 2005). What these previous studies do not provide is a comparative understanding of the social contexts that predict sentiment toward immigrants. Case studies investigating social context have offered detailed analysis on how specific social, economic, and political environments have shaped attitudes to immigrants (Allport, 1954; Portes & Sensenbrenner, 1993; Portes & Zhou, 1993), but they have tended to focus on a few target groups, a few issues per study, and relatively short intervals of time.

The goal of this investigation is to evaluate the psychology of outgroup prejudice through language patterns. Words appearing on popular mass media won the competition of public attention and therefore they provide important clues of collective psychological dynamic, which can lead to inferences of how people are thinking and feeling about a specific group at societal level. For example, it has been found that language from news published on the New York Times informs change of stereotype associated with gender and a few ethnic groups (Garg, Schiebinger, Jurafsky & Zou, 2018). The advantage of identifying language patterns of outgroup prejudice is that it allows a fine-grained comparison of perceptions and social contexts associated with a large number of outgroups.

Construal level theory offers a theoretical foundation to extract quantifiable information relevant to inter-group contact from large corpora. According to construal level theory, the further removed an object is from direct experience, the more abstract the mental representation of that object (Trope & Libermann, 2010). From this perspective, it is fair to predict that people who lack direct experience with an immigrant group will have a more abstract construal of its members. Both laboratory and natural experiments support this prediction. For example, in their analysis of around 700,000 Twitter feeds that mentioned specific social groups, Sneffjella and Kuperman (2015) found that language became less concrete as users moved from describing family to friends to neighbors to coworkers to "foreigners" in general. It has also been shown that people use more concrete language when writing from a first-person than from a third-person perspective (Pronin & Ross, 2006),

indicating that concrete language is more likely to reflect direct contact. Moreover, a recent study on dehumanization of immigrants found that in a task judging punishment for illegal cross-bordering, people who would send immigrants to jail for more time also describe immigrants negatively and in impersonal reference (e.g. it, who) (Markowitz & Slovic, 2020).

Against this background, we investigated perceptions of a wide variety of immigrant groups by analyzing patterns of language surround them. We hypothesized that perceived social distance to an immigrant group is strongly associated with sentiment toward that group. We operationalized social distance in terms of the concreteness of the language used to refer to immigrants in a large corpus of newspaper articles. Specifically, we analyzed language relating to 60 ethnic and religious groups in a corpus containing news articles published in the *New York Times* over a 20-year period (Sandhaus, 2008). We constructed social group corpora by grouping together articles that mention the same ethnic or religious group. Sentiment toward a social group was calculated by computing the mean valence of all words in the respective social corpus, using the valence norms of Hollis, Westbury and Lefsrud (2017). Similarly, we used Hollis et al.'s (2017) concreteness norms to assess the concreteness of the language used to describe each immigrant group (e.g., Hills & Adelman, 2015).

The language concreteness measure allows us to compare perceived social distance to different social groups on the same metric. To go one step further and uncover how each social group were characterized in immigration contexts, we construct an immigration corpora by selecting all articles containing at least one occurrence of the word 'immigrant' or its inflections, and applied Latent Dirichlet Allocation (LDA; Blei, Ng, & Jordan, 2003) to extract topics associated with immigration. LDA is an unsupervised machine learning algorithm that uses Bayesian inference to cluster language based on underlying patterns (or topics) that best explain corpus structure. This approach allowed us to tease apart the underlying social contexts that may drive positive or negative sentiment and to evaluate their impact on change in sentiment toward immigrants over time.

Materials & Methods

Subcorpus of the *New York Times* Annotated Corpus

The *New York Times* Annotated Corpus (Sandhaus, 2008) contains nearly all articles (over 1.8 million) published by the New York Times between January 1981 and June 2007. It can be accessed with a license through Linguistic Data Consortium (<https://catalog.ldc.upenn.edu/LDC2008T19>). We create three sub corpora in this study (see Figure 1 as an illustration). To construct social group corpora for 60 social groups, we collate all articles that contain the same social group (e.g. *Mexican* or *Muslim*). Next, we create an immigration corpus by retrieving all articles that included at least one occurrence of the word ‘immigrant’ or its inflections¹. Lastly, we create a set of social immigrant group corpora that takes the intersection of the above two corpora, that is, articles that mention social groups in the context of immigration (e.g. the Mexican social immigrant corpus contains all articles containing the word ‘Mexican’ and ‘immigration’ or any of combinations of their inflections). The proportion of intersection (articles in social group corpora mentioned in the immigration contexts) ranges from Australian (4%) to Guyanese (57%).

Out of the 60 social groups, 50 were defined by country of origin. Only the largest 50 groups (each more than 0.8% of the total population) reported in the American Community Survey (U.S. Department of Homeland Security, 2010) were included. Following Bogardus (1927) and Parrillo and Donoghue (2005), we also included eight broader ethnic categories (e.g., African American, Muslim, Jew). The authors added two more religious groups, Christian and Buddhist.

Figure 1. The three corpora.

¹ Inflections of *immigrant* includes *immigrants*, *immigration*, *immigrate*, *immigrated*, *immigrating*, etc.

Corpus concreteness and valence

We used the dataset provided by Hollis et al., (2017), which contains valence and concreteness rating for 78,286 English words. This is based on a well-validated computational approach to extrapolate valence and concreteness information from human-rated scores of valence (Warriner et al, 2013) and concreteness (Brysbaert et al., 2014).

Topic modelling

LDA assumes that a set of latent patterns (or topics) explains and generates the structure of textual documents. It computes the distribution of topics over documents, with topics represented as distributions of words. An LDA was trained on the immigrant corpus such that each document was assigned a distribution of topics, and each topic was made up of a distribution of words.² For instance, ‘dangerous illegal workers’ may be translated to ‘10 2 2,’ indicating that the last two words were generated by topic 2 and the first by topic 10. The same word may be assigned to different topics, allowing generic words to appear in multiple topics.

In this study, we set the topic number to 15 to ensure that the model was sufficiently simple (to avoid overfitting) while providing sufficient topic resolution (e.g., to avoid assigning different contents to the same topic). No consensus has yet been reached on a nonarbitrary solution for determining the optimal topic number. In our analysis, the number of topics was chosen to maximize interpretability.

We examined the 10 most relevant words for each topic. We defined the relevance of term w to topic k (Sievert & Shirley, 2014) as:

$$\gamma(w, k|\lambda) = \lambda \log P(w|k) + (1 - \lambda) \log \frac{P(w|k)}{P(w)} \quad (1)$$

where $P(w|k)$ is the probability that term w is assigned to topic k and $P(w)$ is the marginal probability of term w in the corpus. The first component of the equation, $P(w|k)$, prioritizes terms with high frequency in a topic. However, it does not consider how unique term w is to topic k , which can be captured by $\frac{P(w|k)}{P(w)}$, a quantity that Taddy (2011) called *lift*. We set λ to 0.5 to take both components into consideration; λ determines the weight given to the probability of term w under topic k relative to its lift.

Topic specificity

² We used R lda library (Chang, 2012) to train the LDA model for multiple numbers of topics (from 10 to 20) using 1000 iterations. The hyperparameters alpha and beta were set to 0.01 to encourage the model to assign topics to documents such that each document was composed of a few topics and to learn topics that produce a few words with high probability.

We used Equation 2 to compute the specificity of topic k to the immigrant subcorpus as compared with the corpus as a whole:

$$\text{Specificity}(k) = \sum_{i=1}^n \left(\frac{\gamma(w_i|k)}{\sum_{i=1}^n \gamma(w_i|k)} * \frac{P(w_i|\text{immigrant corpus})}{P(w_i|\text{general corpus})} \right) \quad (2)$$

where $\frac{\gamma(w_i|k)}{\sum_{i=1}^n \gamma(w_i|k)}$ is the normalized relevance of word w_i to topic k , and $\frac{P(w_i|\text{immigrant corpus})}{P(w_i|\text{general corpus})}$ is the ratio of the frequency of word w in the immigrant corpus to its frequency in the source corpus. Specificity can range from 0 to near infinity. A specificity of 1 means that, on average, the words characterizing the topic have the same frequency in both the immigrant corpus and the source corpus. Higher topic specificity suggests that they are more likely to occur in the immigrant subcorpus than elsewhere in the corpus.

Topic valence and concreteness

Topic valence and concreteness were computed by averaging the valence and concreteness ratings of the individual words assigned to each topic by LDA.

Distributions of topics over social groups

To determine the strength of association between immigrant topics and ethnic or religious groups, we computed the document-normalized probability distribution of words in social corpora over the 15 topics, with the loading of an social group on topic t being:

$$l_t = \frac{\sum_{d \in D} P_{dt}}{\sum_{t \in T} \sum_{d \in D} P_{dt}}, \quad (3)$$

where d is a document from an social corpus D , t is one of the 15 topics, and P_{dt} is the proportion of words in document d assigned to topic t .

Results

Immigrant Sentiment and Concreteness

Validation. To understand the mechanisms driving sentiment toward immigrants, we computed the valence and concreteness of the language used in reference to each of the 60 social groups. Valence and concreteness were inferred from both the social group corpora and the social immigrant group corpora.

First, we validate our results against Parrillo and Donoghue's (2005) survey on perceived social distance of 30 minority groups in the U.S. The scale they used was the Bogardus social distance scale (Bogardus, 1927) in which participants were asked to evaluate their willingness to take members of the social group in question into their social circles of various degree of intimacy. Social circles range from close relatives and personal friends to foreign visitors. One typical question was "would you be willing to have a member of this group as your colleague at work?"

We found that both concreteness and valence are strongly correlated with Parrillo and Donoghue's survey of social distance (Table 1). Although we argued that concreteness of language is a proxy for perceived social distance, valence has an even stronger correlation with the survey data of social distance than concreteness. This is not entirely surprising because instead of capturing actual inter-personal contact with minority groups, their survey, using hypothetical questions, captures willingness of contact, which reflects a mixture of perceived social distance, affective feelings towards minority groups, and possibly moral considerations.

Table 1. Validation using Parrillo and Donoghue's (2005) survey data on social distance

	Social group corpora	Social immigrant group corpora
Valence x Parrillo	-0.69 ***	-0.72 ***
Concreteness x Parrillo	-0.37*	-0.55**

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Number of social and religious group studied in Parrilo & Donoghue's (2005) study is 30.

Relationship between concreteness and valence. The social group described in the most positive terms was Italian; that described in the least positive terms was Iraqi. The two groups also showed high and low concreteness ratings, respectively. Indeed, across groups,

the language associated with more positively viewed social groups was reliably more concrete, $r(59) = 0.77, p < 0.001, 95\% \text{ CI} = 0.64\text{--}0.86$ (Figure 2).

Overall, the pattern resembles that emerging from Parrillo and Donoghue's (2005) study with American college students: the most positive sentiment was observed for European groups and the least positive for groups from the Middle East, with a variety of Asian groups falling in between.

It can be ruled out as an alternative explanation for these findings that the correlations between valence and concreteness that emerged across social groups might be attributable to linguistic properties of the English language. At the individual word level, there is only weak positive correlation between valence and concreteness across the 13,384 English words in the Warriner et al. (2013) Affective Norms data set; Pearson's $r(13,383) = 0.10, p < 0.001, 95\% \text{ CI} = 0.08\text{--}0.11$. This result is consistent with previous findings that abstract words are less neutral and more emotionally valenced (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011; Vigliocco et al., 2013). At the article level, the correlation between concreteness and valence across the social group corpora ($N = 1,260,046$) was just 0.16 ($p < 0.001, 95\% \text{ CI} = 0.16\text{--}0.17$). The substantial correlations we found across social groups are therefore unlikely to be an artefact of linguistic properties of the English language.

Figure 2. Relationship between valence and concreteness of language in the social group corpora. The size of the dots represents the size of each corpus.

We explored two alternative explanations. The first is media exposure, operationalized in terms of the number of articles mentioning the respective target group. If social contact reduces intergroup prejudice, frequency of exposure to outgroup information could potentially achieve a similar effect. Secondly, a disproportionate emphasis on

immigrant status may be associated with more negative attitudes. For example, the ethnic identify of a terrorist may be more likely to appear in news headlines if the person is an Iraqi than if the person is a European. We operationalize immigrant status as the ratio between the number of articles mentioning a social group in immigrant contexts and all articles mentioning that social group.

We controlled for both of the above factors in two regression models that predicts valence using concreteness. In the first regression model, we include year fixed effects (Table 2, column 1), to control for potential biases generated by shocks common to all minority groups in a given year (e.g. 911 terrorist attack in 2001). We found that the strong positive relationship between concreteness and valence is still robust to the introduction of year fixed effects, media exposure, and immigrant status.

Introducing group fixed effects in the second regression model allows us to explore the relationship between valence and concreteness for each social group over the 20 years (Table 2, column 2). The results suggest that the positive relationship between concreteness and valence is less strong at intra-group level ($\beta = 0.15, 95\% CI = 0.12 - 0.19$). This is probably because the limited time span the corpus covers does not allow large change in public perceptions towards social groups. The large difference between marginal R^2 (variance explained by fixed effects) and conditional R^2 (variance explained by fixed effects and random effects) suggests that the majority of the variance is not explained by intra-group differences. In contrast, document number emerges as a powerful predictor of valence ($\beta = -0.37, 95\% CI = -0.44 - -0.30$), suggesting increased media exposure is associated with negative reporting. Lastly, the coefficient of year ($\beta = 0.03, 95\% CI = 0.03 - 0.04$) suggests that the sentiment towards social groups generally become more positive over time.

Table 2. Language concreteness predicts valence

	Year fixed effect	Group-specific trends
Concreteness	0.65*** (0.61 - 0.69)	0.15*** (0.12 - 0.19)
Document number	0.08 ** (0.03 - 0.13)	-0.37*** (-0.44 - -0.3)
Document ratio (Local/Global)	-0.05 * (-0.10 - 0.00)	-0.04* (-0.09 - 0.00)
Year	N.A	0.03*** (0.03 - 0.04)
<i>Marginal R²</i>	0.43	0.17
<i>Conditional R²</i>	0.44	0.88

Note: The concreteness is a significant predictor of valence in an ordinary least squares estimate. The dependent variable is valence per group per year. Column 1 includes year fixed effect to deal with spurious correlations over time and column 2 includes group-specific trend to deal with spurious correlations across groups. Variables are normalized so that they are all centered at 0 with standard deviation equals to 1. The 95% confidence intervals are included inside the parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

The linguistic measure of valence and concreteness allows us to compare different social groups on the same metric. The price for this simple comparison is the loss of idiosyncratic details about the topics with which each group associates. To resolve this, in the next section, we address the question what are the immigration-related topics and how they associate with each social group.

Immigrant Topics

We applied LDA on the immigration corpus to identify topics associated with cultural representations of immigrant identity that may drive positive and negative sentiment. Table 3 shows the 10 most relevant words in each topic (see Equation 1 in the Methods section for a definition of topic-word relevancy). Keywords for the same topic are strongly associated with each other and are clearly distinguishable from those of other topics. We labelled the topics by summarizing their top 20 keywords (see *Topic* column). The results indicate a wide array of topics, with crime, terrorism, and geopolitical conflict among the most negative topics and museums, music and movies, and restaurants among the most positive. These topics reflect many of the issues commonly associated with both the pros and cons of immigration (Alexander, Brewer, & Herrmann, 1999; Borjas, 1990; Carens, 2013; Cuddy et al., 2007; Peabody, 1985; Poppe, 2001; Skeldon, 2014).

Table 3

Top 10 Keywords for Each Topic (From the Most Negative to the Most Positive)

Index	Topic	Keywords
1	Crime	Police, officer, arrest, charge, prosecutor, drug, kill, gang, crime, shoot.
2	Terrorism	Muslim, terrorist, bomb, attack, intelligence, Islamic, FBI, mosque, Sept, Iraq
3	Legal	Immigration, law, court, alien, judge, legal, justice, case, federal, lawyer
4	Politics	Republican, bush, democrat, bill, president, vote, senate, senator, campaign, Clinton
5	Geopolitical conflict	Israel, minister, soviet, France, Germany, Europe, party, prime, Palestinian, Jew
6	Refugees	Refugee, Cuban, asylum, Haitian, unite, Miami, boat, Castro, state, official

7	Illegal employment	Worker, border, Mexican, company, labor, job, wage, work, pay, illegal
8	Census	Hispanic, population, percent, Asian, black, census, Chinese, Korean, Latino, immigrant
9	Neighborhood	City, build, house, neighborhood, county, resident, island, apartment, rent, community
10	Books	Write, book, life, American, world, think, history, story, time, way
11	Religion	Church, catholic, Irish, bishop, priest, Jewish, religious, parish, Pope
12	Education	School, student, child, teacher, education, parent, program, health, care, college
13	Restaurant ³	Restaurant, cook, eat, chicken, room, shop, soccer, dish, food, cup
14	Music & Movies	Theater, film, music, movie, play, art, direct, musical, dance, song, artist
15	Museums	Museum, Sunday, tour, street, information, tomorrow, admission, exhibition, park, sponsor

Note: we combined inflections (e.g. German, Germany) avoid unnecessary duplications. An interaction visualization (generated by R package “LDAvis”) of topic-word association with varying degree of lambda can be accessed at <https://osf.io/tvkn3/>. Note that LDAvis orders topics by topic size, not by valence.

Next, we analyzed linguistic features of these topics: their valence, concreteness, and topic specificity (Figure 3). Topic valence and concreteness is computed by the average of all words assigned to the given topic. We find no significant correlation between topic valence and concreteness ($r(13) = 0.36, p = 0.17$). This indicates that topic association is not the driver between concreteness and valence. Some topics that are more concrete (like crime and terrorism) are not high valence, and the same is true of some high valence topics (like books). Thus the relationship between concreteness and valence for different social groups found in Figure 2 is not the result of language distributed across multiple topics. More specifically, what this tells us is that if a group is associated with a concrete negative topic, like crime, they tend to also be associated with less concrete topics. Negative discussion of social groups and less concrete language tend to go hand-in-hand.

³ λ (of equation 1) were set to 0.3 when displaying key works for topic 13 because this topic is mixed with generic linguistic patterns underlying all articles, such as *say, like, one, day, get, come*. Reducing λ further penalize the weight of high frequency words that tend to appear across all articles.

Figure 3. Concreteness and valence of the 15 immigrant topics identified using LDA. Dot size corresponds to the number of words assigned to each topic. Dot color represents topic specificity, with higher values indicating that the topic was more likely to occur in the immigrant corpus than elsewhere in the New York Times corpus.

It worth to point out that these topics associates with immigration at various degree. It is possible that the more closely a topic associates with immigration, the more abstract and negative its language becomes. The association with immigration can be quantified as topic specificity (see Equation 2). Higher topic specificity suggests that words were more likely to occur in immigration-related articles than elsewhere. We found that topic specificity was negatively correlated with concreteness (Pearsons' $r(13) = -0.59$, $p < 0.05$, 95% CI = $-0.85 - -0.11$) and valence ($r(13) = -0.60$, $p < 0.05$, 95% CI = $-0.85 - -0.13$). In other words, topics more specifically associated with immigration tended to be more abstract and negative.

Figure 4: Distributions of topics over social groups ranked by valence. The topics identified in Table 1 are plotted on the x-axis. The y-axis shows the normalized weighting of each topic on each social group. Topics are arranged by valence, with the lowest (in red) on the left and

the highest (in green) on the right; social groups are also ranked by overall valence, with the most negative in the top left corner and the most positive in the bottom right.

To understand which topics defined the media representation of each social group in the context of immigration, we computed the document-normalized probability distribution of words in social immigrant group corpora over the 15 topics (see Equation 3). Figure 4 presents the distributions of topics over social groups. Groups ranking lower on mean sentiment were associated primarily with negative topics. For example, the groups Iraqi, Palestinians, Lebanese, and Syrian were represented mostly in contexts of terrorism or geopolitical conflicts; the groups Cuban, Nicaraguan, Vietnamese, and Venezuelan were closely associated with the topic refugee; Mexicans were mentioned primarily in the context of illegal workers. In contrast, groups with high sentiment rankings were closely associated with positive, less immigrant-specific topics (e.g., restaurants, museums, movies and music)—and rarely represented in the context of negative topics.

Figure 5. Using the distribution of words over the 15 topics to predict the valence of immigrant groups: Regression coefficients from an averaged linear regression model. Error bars represent the 95% confidence interval.

To identify the topics driving sentiments toward immigrants, we regressed the valence of social groups (inferred from the social group corpora) on the distribution of words over the 15 topics as input. As the model contained 15 independent variables and just 60 data points, we used elastic net regularization, a combination of Lasso regression and ridge regression. These techniques perform simple linear least squares regression, but penalize the coefficients

of the inputs x_i based on their size. The penalty forces some regression coefficients to zero. We cross-validated our findings by dividing our data set into 10 equal groups, training our model on a random sample of seven groups, and predicting immigrant sentiment in the remaining three. This cross-validation exercise was repeated 1,000 times to calculate the average adjusted R^2 for the out-of-sample predictions and average regression coefficients. 78% of the variance in sentiment toward immigrant groups can be explained by topic profiles of individual immigrant groups. Overall, the negative topics had a stronger impact on sentiment than the positive topics did (Figure 5). Three major negative topics—crime, terrorism, and legal—significantly predicted negative sentiment toward immigrants. Restaurant was the topic most strongly predicting positive sentiment. Politics, geopolitical conflict, refugees, illegal workers, and religion did not significantly predict sentiment toward immigrants.

Discussion

Despite much research on prejudice and stereotypes, large-scale quantitative studies of opinion change towards multiple immigrant groups are scarce. This study makes a number of contributions to this area of research. First, we found that perceived social proximity to an immigrant group, operationalized in terms of the concreteness of the language used in reference to that group, was strongly correlated with positive sentiment. We further identified topics associated with each immigrant group and showed how those topics were related to change in sentiment toward 60 immigrant groups.

Our approach maps onto two essential components of intergroup contact theory: frequency and quality of contacts (Allport, 1954; Brown & Hewstone, 2005). A more concrete representation of a social group may reflect a higher frequency of contact: more frequent social contacts afford more specific, idiosyncratic, and incidental information. However, as our result shows, increased exposure alone is not sufficient. In fact, our analysis on the New York Times corpus shows that increased exposure associates with even more negative reporting. Topic model approach complements the quantitative analysis of frequency of social contacts by offering a more microscopic, qualitative examination of the content of those contacts. Our finding that use of concrete language predicts a more positive sentiment paints an optimistic picture—one in which encounters in positive environments may improve perceptions of immigrant groups. The topics we identify here therefore represent specific predictions for future research about what kinds of associations may drive positive or negative sentiment.

We have also shown that immigrant groups can be represented along a continuous measure of social distance. Classic theories on outgroup negativity have focused on an ingroup vs. outgroup dichotomy, thus overlooking differences among outgroups – a cognitive bias that these prejudice theories have identified as one of the symptoms of outgroup bias. In contrast, the more recent works from Fiske, Cuddy, Glick, Xu (2002) highlights the differences among stereotypes, proposing that outgroups are systematically perceived along two dimensions of warmth and competence (the stereotype content model). By offering a continuous measure, we complement the previous research on variance of stereotype contents by explaining which outgroups are more psychologically distant than others and how such differences relate to perceived sentiment and cultural representations.

The fact that our findings on social distance are largely consistent with the survey results of Parrillo and Donoghue (2005) suggests that our corpus approach captures meaningful patterns despite its possible limitations. As the second largest news distributor in the U.S., with its headquarters in a metropolitan city, the *New York Times* is well-positioned to offer wide coverage of immigration issues and to influence its readers' attitudes toward outgroups. Nonetheless, the *New York Times* is unlikely to be representative of the full diversity of public opinion, as it may be experienced in, for example, online and offline social communities. We further acknowledge that the topics identified may differ across media targeting different audiences. However, given the established theory on which we frame our approach, the relationship we found between social distance and sentiment represents a hypothesis about the topics that may be found in other contexts, such as everyday conversations or communication on social networks such as Twitter. Unlike the *New York Times*, these channels are not required by style and editorial rules to use formal and politically correct language. We can therefore expect socially distant outgroups to be associated with even more negative language in these channels. One direction for future research is to analyze language produced by different social communities.

Overall, we believe that the strengths of the corpora approach outweigh its limitations. These strengths include ecological validity, achieved by (a) evaluating outgroup prejudice through language patterns, (b) tracking relationships between social distance and sentiment in a data set referencing a large variety of social groups, (c) studying perceptions of immigrants outside the laboratory, thus avoiding problems such as socially desirable response bias and (d) providing a valid social scientific approach with specific language patterns to potentially flag certain outgroups at greater risk of prejudice.

References

- Alexander, M. G., Brewer, M. B., & Herrmann, R. K. (1999). Attitudes and affect: A functional analysis of out-group stereotypes. *Journal of Personality and Social Psychology*, 77(1), 78–93.
- Allport, G. W. (1954). *The nature of prejudice*. New York, NY: Addison.
- Ames, D. R. (2004). Inside the mind reader's tool kit: Projection and stereotyping in mental state inference. *Journal of Personality and Social Psychology*, 87(3), 340–353.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993–1022.
- Bogardus, E. S. (1927). Race friendliness and social distance. *Journal of Applied Sociology*, 11, 272–287.
- Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett, *Advances in Neural Information Processing Systems 29* (pp. 4349–4357). Retrieved from <http://papers.nips.cc/paper/6228-man-is-to-computer-programmer-as-woman-is-to-homemaker-debiasing-word-embeddings.pdf>
- Borjas, G. J. (1990). *Friends or strangers: The impact of immigrants on the US economy*. New York, NY: Basic Books.
- Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings* (Tech. Rep. No. C-1). Gainesville, FL: University of Florida, The Center for Research in Psychophysiology.
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, 86(2), 307–324.
- Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology*, 37(37), 255–343.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904–911.
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186.
- Carens, J. (2013). *The ethics of immigration*. New York, NY: Oxford University Press.
- Chang, J. (2012). Collapsed Gibbs sampling methods for topic models. R Package.

- Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631–648.
- Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., ... & Htun, T. T. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, 48(1), 1–33.
- Fennelly, K., & Federico, C. (2008). Rural residence as a determinant of attitudes toward US immigration policy. *International Migration*, 46(1), 151–190.
- Garg, Nikhil, et al. "Word embeddings quantify 100 years of gender and ethnic stereotypes." *Proceedings of the National Academy of Sciences* 115.16 (2018): E3635-E3644.
- Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12), 517–523.
- Hills, T. T., & Adelman, J. S. (2015). Recent evolution of learnability in American English from 1800 to 2000. *Cognition*, 143, 87–92.
- Hollis, G., Westbury, C., & Lefsrud, L. (2017). Extrapolating human judgments from skip-gram vector representations of word meaning. *The Quarterly Journal of Experimental Psychology*, 70(8), 1603-1619.
- Huntington, S. P. (1993). The clash of civilizations? *Foreign Affairs*, 72(3), 22–49.
- Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.
- Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140(1), 14–34.
- Markowitz, D. M., & Slovic, P. (2020). Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences*, 117(17), 9260-9269.
- Paluck, E. L., Green, S. A., & Green, D. P. (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy*, 3(2), 129-158.
- Parrillo, V. N., & Donoghue, C. (2005). Updating the Bogardus social distance studies: A new national survey. *The Social Science Journal*, 42(2), 257–271.

- Peabody, D. (1985). *National characteristics*. New York, NY: Cambridge University Press.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, 5(4), 461–476.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783.
- Poppe, E. (2001). Effects of changes in GNP and perceived group characteristics on national and ethnic stereotypes in Central and Eastern Europe. *Journal of Applied Social Psychology*, 31(8), 1689–1708.
- Portes, A., & Sensenbrenner, J. (1993). Embeddedness and immigration: Notes on the social determinants of economic action. *The American Journal of Sociology*, 98(6), 1320–1350.
- Portes, A., & Zhou, M. (1993). The new second generation: Segmented assimilation and its variants. *The Annals of the American Academy of Political and Social Science*, 530(1), 74–96.
- Pronin, E., & Ross, L. (2006). Temporal differences in trait self ascription: When the self is seen as an other. *Journal of Personality and Social Psychology*, 90, 197–209.
- Sandhaus, E. (2008). The *New York Times* Annotated Corpus. Philadelphia, PA: Linguistic Data Consortium. Retrieved from <https://catalog.ldc.upenn.edu/LDC2008T19>
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robbers Cave experiment*. Norman, OK: University Book Exchange.
- Sievert, C., & Shirley, K. E. (2014). LDAvis: A method for visualizing and interpreting topics. In J. Chuang, S. Green, M. Hearst, J. Heer, & P. Koehn (Ed.). *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces* (pp. 63–70). Baltimore, MD: Association for Computational Linguistics.
- Skeldon, R. (2014). *Migration and development: A global perspective*. New York, NY: Routledge.
- Snefjella, B., & Kuperman, V. (2015). Concreteness and psychological distance in natural language use. *Psychological Science*, 26(9), 1449–1460.
- Taddy, M. (2012). On estimation and selection for topic models. In N. D. Lawrence & M. Girolami, *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)* (pp. 1184–1193). Retrieved from <http://proceedings.mlr.press/v22/taddy12/taddy12.pdf>

- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33(1), 1–39.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–463.
- The World Bank. (2012). *International migrant stock*. Retrieved from https://data.worldbank.org/indicator/SM.POP.TOTL?end=2017&name_desc=false&start=1962&view=chart
- U.S. Department of Homeland Security. (2010). *Yearbook of immigration statistics 2017*. Retrieved from: <https://www.dhs.gov/immigration-statistics>.
- Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2013). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, 24(7), 1767–1777.
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191–1207.