

Measuring racial bias within the Dutch public
news outlet's coverage

Master Thesis
Applied Data Science

Dimitri de Boer

0379689

Wednesday 13th July, 2022

Utrecht University

Supervisor: Anastasia Giachanou

Second Examiner: Ayoub Bagheri



Abstract

This study provides an analysis of portrayals by the Dutch public news outlet of 6 of the most populous non-western ethnicity groups in the Netherlands. A dataset is used that contains the whole collection of news articles published between January 1st 2010 and May 1st 2022.

Two methods are employed to compare outgroup members (non-western ethnic groups) with the native group (the Netherlands) and ingroup members (western ethnic groups). First, a continuous-bag-of-words embedding model is trained on the entire collection of articles. Herein, associations between stereotypes and ethnic groups are measured by capturing the distances between ethnic groups and stereotype indicators that signal hostility, deviance, threatening behavior or objects, criminality, judicial authorities, and/or illegal activities (high-threat), and (un)intelligence, low education levels, unemployment, addiction, and/or homelessness (low-status). Second, the National Research Council Canada (NRC) lexicon is applied for its sentiment and polarity scores. Articles that discuss different ethnic groups are compared for their sentiment outputs to identify how different ethnic groups are discussed in comparison to the other.

The results of this study show ethnic outgroups to be closer in proximity to low-status stereotypes in comparison with the native group and ingroups. This was not found for high-threat stereotypes. In addition, the sentiment analysis revealed articles that discuss ethnic outgroups to have a more negative tone, more expressed negative emotion and less expressed positive emotion. The findings indicate racial bias might be present within the Dutch public news outlet's coverage.

Contents

Abstract	1
1 Introduction	3
2 Literature Review	6
2.1 Computationally Identifying Ethnic Bias	6
2.2 Sentiment Analysis	8
3 Data	8
3.1 Dataset	9
3.2 Pre-Processing	9
3.3 Word Embedding	9
3.4 Indicator Lists	10
3.5 NRC Lexicon Sentiment Scores	10
3.6 Retrieve Ethnic Names and Ethnic Subsets	11
4 Method	11
4.1 Stereotype Retrieval	12
4.2 Sentiment Analysis	12
5 Results	13
5.1 Stereotype Proximities	13
5.2 Sentiment Investigation	17
5.3 Discussion on Results	22
6 Conclusion	24
7 Limitations and Future Work	24
References	25
Appendix: High-Threat and Low-Status Indicator Lists	a

1 Introduction

Bias is a prejudice that can be beneficial or harmful for a person or group in comparison with another. Mostly the biased prejudice is perceived as unfair and therefore unwanted. There are different types of biases. Racial bias is the differential treatment of people of colour because they belong to a certain race (Small & Pager, 2020). The stereotypical associations that are linked to a person's race can be harmful because they contribute to the formation and manifestation of negative implicit attitudes, which result in discriminatory behavior (Arendt & Northup, 2015). Furthermore, when media outlets, journalists, and reporters show bias in what stories to report and what not, and in what way stories are reported, this is called media bias (Bharathi & Geetha, 2019). A media study investigation on immigrant news within different US news outlets found that these articles are overall skewed towards a negative sentiment (Van Klinger, Boomgaarden, Vliegthart, & De Vreese, 2015). News media have been linking immigrant streams to problems such as increasing crime rates, higher threat of terrorism, and the shrinking of the socio-economic state even though there seems to exist little relationship between real-world events and immigration news (Jacobs, Damstra, Boukes, & De Swert, 2018). Furthermore, it is investigated that inaccurate reporting in media can promote misconceptions towards the reported subject (Garrett, Weeks, & Neo, 2016). Similarly, it is investigated that peoples' beliefs about the political world differ based on which media platform is used for information consumption Meirick (2013). The media influence can have a big impact on shaping or changing the opinions of people as well as the attitude one can have towards others. Therefore, ethnic groups need to be represented accurately and media bias should be eliminated. One way to capture the problem is to analyse the existing media texts that can help us to discover meanings, purposes and sentiments expressed in them (Guo & Zhang, 2020).

The terms ingroup and outgroup originate from Tajfel, Billig, Bundy, and Flament (1971). The terms were introduced as a way of describing behaviors between groups. The distinctions for in- and outgroups were made to refer to members who are perceived to belong to the same group as themselves and to refer to members who are perceived to be different from oneself on the basis of any possible social criteria that can draw a line between 'them' and 'us'. In this paper the terms ingroup and outgroup are used to make a distinction between western ethnicities and non-western ethnicities.

The study topic of prejudice based on race is relevant to this date. A meta-analysis, collecting data from 43 audit studies within OECD (The Organization for Economic Cooperation and Development) countries, in the period between 1990-2015, found that ethnic outgroups had to apply for jobs about 50% more than the ingroups of equally qualified applicants. This finding was valid even for countries that have anti-discrimination laws (Zschirnt & Ruedin, 2016). Moreover, the Dutch government recently plead guilty to institutional racism claims regarding the tax authorities. The tax authorities included minority groups far more easily in high-risk groups. As a result, the application of these people were

rejected when applying for debt restructuring programs ¹. This shows unfavorable discrimination based on ethnicity is still present and should be an alert that a lot of research should still be done, even where one may believe it is to be fair already.

Besides, the prominent public opinion about immigrants is seen as unfavorable in most European societies (Gorodzeisky, 2013). Sociologists have pointed out that due to the arrival of minorities through immigration, negative sentiments towards these minorities have risen in the past decades (Gorodzeisky & Semyonov, 2016). The perception of extra competition over existing socio-economic resources together with safety threats have made these negative sentiments worse. The prominent public opinion is that the presence of immigrants may lead to conflicts about the resources and an increase in crime and terrorism (Kentmen-Cin & Erisen, 2017).

Within the literature, several methods for identifying racial bias with natural language processing exist. One powerful tool to analyze texts is with the use of word embeddings. In these models, words are vectorized in k length vectors representing a location in k dimensional vector space. Herein, similar words have locations nearby each other. Embeddings are known to capture more semantics than simple co-occurrence analysis. For example, word embedding models can capture that PlayStation is similar to Xbox and Italy is similar to France and Austria (Caliskan, Bryson, & Narayanan, 2017; Collobert et al., 2011).

Another advantage of word embeddings is that they can capture word relationships. Distances between words can also capture meaning in such a way that in multidimensional vector space, the distance between London and England is equal to the distance between Paris and France. The word embedding model is also shown to be effective in capturing stereotypes within large corpora of texts (Bolukbasi, Chang, Zou, Saligrama, & Kalai, 2016). The stereotype is learned by the embedding model and is reflected in the vector space. For example, a typical bias in the word embeddings is that adjectives that represent strength and leadership can be closer to man whereas adjectives more related to emotions are closer to woman. In summary, the embedding model can capture semantic meaning derived from the corpus of text. When relations between words are different from expected, this can signal bias in the data.

In addition to word embeddings, sentiment analysis is applied to identify the tone, as well as the emotions expressed within texts where ethnic groups are discussed. Sentiment analysis is the extraction of expressed private states, opinions and affect of the speaker towards a target entity in texts (Wiebe, 1994). Sentiment analysis can be performed automatic with the help of computational sentiment analyzers that can output scores for polarity and specific emotions. Polarity is a positive/ negative score and emotion scores often scope to a set of basic emotions. Inspecting the tone and emotion of articles that discuss different ethnic groups will aid in identifying the tone and attitude that is expressed when people from these ethnic groups are discussed in news articles.

¹<https://nos.nl/artikel/2430805-kabinet-erkent-institutioneel-racisme-bij-deel-fiscus-heeft-velen-pijn-gedaan>

This paper focuses on six of the most populous non-western ethnic groups in the Netherlands. *Turkish, Syrian, Moroccan, Somali, Afghan, and Iraqi* are among these ethnic groups. The ethnic groups will be grouped together as outgroup and investigated in relation to indicator words that signal high-threat and low-status. The outgroup will be compared with ingroup members and the native group to identify relative differences. This method is similar to the applied method in the paper by Kroon, Trilling, and Raats (2021). Fiske’s stereotype content model is employed to identify stereotypes in texts. The model states two dimensions a person or group can be predisposed in (Fiske, Cuddy, Glick, & Xu, 2018). According to the stereotype content model, outgroups are frequently classified in two mixed dimensions: paternalized groups that are perceived as warm but also disrespected as incompetent (low-status), and groups that are admired for their competence but disliked due to a lack of warmth (high-threat). The category into which an ethnic group falls is determined by structural relationships between the ethnic and native groups, specifically status and competition.

This paper investigates racial bias present in the Dutch public news outlet’s coverage, the NOS. The focus will be on the following two research questions:

RQ 1 *How do stereotypes differ in non-western ethnicities compared to western and native ethnicities?*

RQ 2 *How do sentiment distributions differ within articles that contain non-western ethnic names compared to articles that contain native and western ethnic names?*

News media have linked non-western ethnic outgroups as a threat to socio-economic resources and safety through crime and terrorism linkage as well as practicing jobs of low economic status (Jacobs et al., 2018). The hypotheses therefore state a closer link in the word embedding between non-western ethnicities and the stereotype associations. The first hypothesis is that the Dutch public news outlet implicitly associates non-western ethnicities closer with low-status indicators in comparison with western and native ethnicities. The second hypothesis is that the Dutch public news outlet implicitly associates non-western ethnicities closer with high-threat indicators in comparison with western and native ethnicities. With regards to the first research question, two hypotheses are formulated as follows: outgroups will be more associated to low-status stereotypes in comparison with native and ingroups (H1) and outgroups will be more associated with high-threat stereotypes in comparison with native and ingroups (H2).

The discussion of non-western ethnic groups will be further investigated with the use of a sentiment analysis tool. It is hypothesized that the overall presence of ethnic people in news coverage will be more negatively toned compared to native and ingroups. The third hypothesis is as follows: articles that contain

non-western ethnic names have a more negative tone, more expressed negative emotion and less expressed positive emotion than articles that contain native- and western ethnicity names. With regards to the second research question, the third hypothesis states that the sentiment distribution for outgroups will display more negative emotions, less positive emotions and a negatively skewed polarity score compared to native and ingroups. To summarise, we have the following hypotheses:

- H1** The Dutch public news outlet implicitly associates non-western ethnicities closer with low-status indicators in comparison with native and ingroups.

- H2** The Dutch public news outlet implicitly associates non-western ethnicities closer with high-threat indicators in comparison with native and ingroups.

- H3** articles that contain non-western ethnic names have a more negative tone, more expressed negative emotion and less expressed positive emotion than articles that contain native- and western ethnicity names.

2 Literature Review

In this section, the existing literature regarding embedding models and sentiment analysis for stereotype- and media coverage investigation is discussed.

2.1 Computationally Identifying Ethnic Bias

Recently, there has been an increase in the computational identification of social biases as a result of an increase of the efficacy of identifying these biases. Previously, literature on media biases was limited due to the scarcity of accurate measures (Arendt & Karadas, 2017). Prior literature on the identification of ethnic bias applied word co-occurrence methods that identify word pairs within sentences or within articles of specific target words and attributes (gender references and occupations/ ethnic references and terrorism) (Rekabsaz, West, Henderson, & Hanbury, 2021; Ruigrok & Van Attevelde, 2007). The goal of co-occurrence methods is to calculate associations between concepts. An association is defined by the frequency of the concepts co-occurring together within texts or sentences (Van Attevelde, 2008). A limitation of co-occurrence methods is that they cannot capture contextual semantic information, unless it is applied

manually. However, a manual search is very time consuming when analyzing large collections of text. An example of an applied manual co-occurrence method is the research of Kroon, Van Selm, Ter Hoeven, and Vliegthart (2018). The research identified stereotypes regarding older employees by manually coding news articles that discuss the target group and whether different stereotype categories are present. The research found a negative stereotype association with the target group and competence. Stereotypes in the news media can be subtle and therefore difficult to detect computationally. Modelling the text with use of the word embedding model has made it easier to capture social biases in text corpora as the model can very powerfully identify biases and manual inspections are spared (Bolukbasi et al., 2016; Caliskan et al., 2017; Garg, Schiebinger, Jurafsky, & Zou, 2018). The advent of word embedding models has led to a relatively new field, the one of detecting biases in text corpora computationally.

Word embedding models have been used to identify ethnic bias in several Dutch newspapers' articles by Kroon et al. (2021). The researchers found stereotype dimensions of non-western outgroup members to be reflected closer to stereotype dimensions of high-threat and low-status within the trained embedding model. These dimensions arise from the Stereotype Content Model proposed by Fiske et al. (2018). The stereotype dimensions signal hostility, deviance, threatening behavior or objects, criminality, judicial authorities, and/or illegal activities (high-threat), and low-social class, (un)intelligence, low education levels, unemployment, addiction, and/or homelessness (low-status). The list of indicators are created by manual selection of indicator words that occurred in the corpus of texts. The result indicates that the newspapers' articles represent outgroups to be implicitly more associated with high-threat and low-status stereotypes and therefore indicate ethnic bias to be present within the analyzed corpus of text.

In addition, the research of Sorato, Zavala-Rojas, and Ventura (2021) did a similar investigation to the article of Kroon et al. (2021). The research also investigated newspaper articles, now from a Spanish newspaper, and investigated the trained embedding model's captured distances to drug use, prostitution, crimes and poverty. The ethnic groups' distances towards the indicators identify implicit stereotype associations of different nationalities. In addition, the researchers compared the calculated distances with the Gross Domestic Product of the country of origin. The indicator lists were derived from the same list the research of Kroon et al. (2021) had applied, but were separated in different categories. This study found that the articles in the newspaper represents Colombian, Ecuadorian, Moroccan and Romanian outgroups to be implicitly more associated with the stereotype indicators compared to native and ingroups.

Next to the investigation of ethnic bias with the use of word embedding models, the research of Wevers (2019) investigated the presence and evolution of gender bias within and between several Dutch newspapers' articles from the period 1950-1990. The research found clear changes within newspapers over time as well as differences between newspapers. The applied method to identify these biases also included the investigation of differences between specific references and attributes within a trained word embedding model.

2.2 Sentiment Analysis

Sentiment analysis tools have been used for several purposes such as identifying sentiments towards travel destinations from twitter texts regarding tourist arrivals (Starosta, Budz, & Krutwig, 2019), mapping the sentiments of touristic reviews (Chaabani, Toujani, & Akaichi, 2018) and measuring sentiments towards specific topics within articles (Ren & Hong, 2017).

To identify the sentiment in the discussion of different target groups, sentiment analysis has recently been used to investigate media coverage of Muslims in American newspapers. The researchers Bleich and van der Veen (2021) used a lexicon based method to measure polarity among articles that discuss Muslims and compare them to articles where Hindus, Jews, and Catholics are discussed. Researchers found that the tone is significantly more negative when Muslims are discussed compared to the average newspaper article as well as when Hindus, Jews, and Catholics are discussed. The researchers state that even when controlling for negative events causing the negative tone in articles, the tone is still skewed more towards the negative. The results indicated media bias to be present as journalists can choose what stories to cover, for how long and how extensively it will be covered.

In addition, Guo and Zhang (2020) have investigated sentiment orientation towards eight countries within the international newspaper 'The Economist'. The Watson Natural Language Understanding API, based on deep learning, was utilized to extract sentiment scores. Ingroup countries are compared with the outgroup, China. Ingroup countries consist of the G7, a group of countries that are amongst the worlds' most developed economies which consist of the United States, the United Kingdom, Canada, France, Germany, Italy and Japan. In contrast with aforementioned literature, the study found no bias in media representation. The results showed that articles are mainly oriented towards negative tone. Also they showed that general attitudes towards the ingroup members are not different from attitudes toward the outgroup member.

This study will build upon the work of Kroon et al. (2021), Sorato et al. (2021) and Wevers (2019) with the aim to investigate ethnic bias in newspaper text by examining a trained word embedding models' captured distances between different ethnic groups. The indicators will, similarly to the article of Kroon et al. (2021), be representing high-threat and low-status as stereotype dimensions. In addition, as is performed by Bleich and van der Veen (2021) and Guo and Zhang (2020), articles' sentiments' will be analyzed and compared between natives, ingroups and outgroups.

3 Data

This section describes the original dataset and all the steps that are taken in the process of data transformation.

3.1 Dataset

The original dataset of this research is publicly available on Kaggle ². The dataset contains all published articles from the Dutch public news outlet, the NOS, from January 1st 2010 until May 1st 2022. To the best of our knowledge, no ethical concerns exist regarding the use of this dataset. The dataset originates from scraping the website of the NOS ³. The website and its content are publicly available, therefore it is believed to be available for anyone who wishes to analyze its contents. The dataset includes information of date and time of publication, the title and content, the original url and the category under which the NOS filed the articles. The NOS also publishes live blogs. These are not included in the data set. The dataset contains 239373 articles with an average article length of 250.29 words.

3.2 Pre-Processing

Pre-processing is necessary to prepare the corpus for the creation of the embedding model. The Spacy package allowed for an efficient way to apply the pre-processing steps ⁴. Lowercasing is applied so the model recognizes words with different capitalization as the same word and compound words that are connected with a '-' are connected so the model understands the word as a single token. In addition, stop words are removed as they do not contribute to the embedding model's information and punctuation is removed so that words that include punctuation are perceived as the same word as to when they do not (at the end of a sentence for example). Also, the removal of punctuation removes tokens that exist of single punctuation, these are seen as noise. Finally, tokenization is performed to transform every word into a single token.

3.3 Word Embedding

After tokenization, vectorization is applied. vectorization transforms the word token into a numerical sequence representation. The embedding model allows for words to be transformed into a sequence of numbers by training on a large collection of text. Words that have similar semantic meanings will have locations nearby in the multidimensional vector space. The applied model is a continuous-bag-of-words model, also known as CBOW. In a CBOW model, a target word is learned from looking at its neighboring words. In this way, the model learns the contexts of the words as well. The model is trained on the whole corpus of NOS articles and is within each article trained to predict target words by looking at its 10 direct neighboring words. Token-representing vectors are of length 100.

²<https://www.kaggle.com/datasets/maxscheijen/dutch-news-articles>

³<https://nos.nl>

⁴spacy.io

3.4 Indicator Lists

Two lists of indicator words are used to signal high-threat and low-status stereotypes, similar to those used in the article by Kroon et al. (2021). The researchers' investigation of ethnic associations with high-threats and low-status dimensions yielded the indicator lists. These are created using a bottom-up approach that involved inspecting and labeling the closest words in vector space to ethnic nouns. The 100 most similar words to each ethnic noun were retrieved and revised in the word embedding. The high-threat list contains words that refer to hostility, deviance, threatening behavior or objects, criminality, judicial authorities, and/or illegal activities. Words referring to low-social class, (un)intelligence, low education levels, unemployment, addiction, and/or homelessness were defined as low-status indicators. The indicator lists mirror real-world dimensions: criminality rates (high-threat) and the reception of social benefits (low-status). The same indicator lists are applied here. The appendix 7 provides an overview of the Dutch high-threat and low-status indicators together with their English translation.

3.5 NRC Lexicon Sentiment Scores

The National Research Council Canada (NRC) lexicon is an open-source word-emotion association lexicon (Mohammad & Turney, 2013). A lexicon-based approach is appropriate in this research case because it requires no training data. It is crowd-sourced annotated and the emotions are based on Plutchik's eight basic emotions; joy, sadness, anger, fear, disgust, trust, anticipation, and surprise (Plutchik, 1980). The choice for representing Plutchik's set of emotions in the lexicon was due to a well-founded coverage in psychological, physiological, and empirical research, a well-balanced emotion set between positive and negative emotions, and Plutchik's basic emotion set is a superset of some of the proposed other sets of basic emotions (Mohammad & Turney, 2013). The lexicon returns per word in their dictionary a combination of emotions the word is associated with as well as a polarity score of either positive or negative. For instance, the word 'terrible' is associated with the emotions: anger, disgust, fear, and sadness, and returns a negative score. Scores are either 0 (not associated) or 1 (associated). The NRC lexicon provides translations to their original English word set. The Dutch word set can then be applied to the pre-processed articles in the NOS data set. As a result of using the same word translation for different English words, some Dutch translations appear multiple times in the lexicon. Because of this, a word may yield more than one set of associated emotions. To address this, the various sets of emotion associations were averaged. Thus, an emotion score does not have to be either 0 or 1, but can also be a float. The emotion scores, as well as the polarity score, are normalized per article resulting in a decimal score between 0 and 1 per emotion per article (all emotion scores sum up to 1 and positive and negative score sum up to 1).

3.6 Retrieve Ethnic Names and Ethnic Subsets

As ethnic names signal the discussion of an ethnic group member. Ethnic name lists are necessary to indicate the relation between the ethnic group and the stereotype dimensions. Forebears' website is scraped to get a list of the 100 most popular names per country. This is done for the outgroup members, ingroup members and the native group's country. For an overview of the most popular names consult Forebears' website⁵. Forebears is a genealogical portal that catalogs on- and offline genealogical sources to make them easily accessible by researchers who are looking for records that belong to specific places; towns, regions, or countries. The website collects per place, the most popular names and ranks them accordingly. The website is scraped for the most popular forenames of 6 of the largest non-western ethnicities present in the Netherlands as well as Dutch native forenames, and Italian and Greek forenames to serve as Western ingroups for comparison. The list of forenames was manually inspected and forenames that were equal to a word occurring in the Dutch language were removed from the lists to prevent including articles in the analysis that do not contain an ethnic name but included the equal word. In total 8 words are removed, these include 'adel', 'ben', 'la', 'lul', 'mona', 'margarita', 'ton' and 'van'. All articles are collected that contain any of the ethnic names and organised to their belonging group to obtain a collection of articles that discuss the native group, out- and ingroups. The number of articles per ethnic subset can be found in table 1. The table shows the native group to have the largest subset of articles, followed by the outgroup and the ingroup, respectively.

Table 1: Sizes of ethnic subsets

Number of articles per ethnic group	n	%
Native	57866	62.4
Outgroup	26617	28.7
Ingroup	8253	8.9
Total	92716	100

4 Method

This section explains the methods for investigating the research questions: *How do stereotypes differ in non-western ethnicities compared to western and native ethnicities?* and *What is the sentiment distribution within articles that contain non-western ethnic names compared to articles that contain native and western ethnic names?*. Included non-western ethnicities are; *Iran, Morocco, Turkey, Syria, Somalia* and *Afghanistan*, and included western ethnicities are; *Greece* and *Italy*. These are also called outgroups and ingroups respectively. The native group is the news outlet's originated country, the Netherlands. Before

⁵<https://forebears.io>

the application of the methods, the corpus of text is pre-processed and vectorized with the trained embedding model as described in the data section. The word embedding model is trained on the entire collection of NOS articles.

4.1 Stereotype Retrieval

For the first research question, the distances between the most popular ethnic forenames and stereotype dimensions high-threat and low-status are captured in the embedding model and compared among groups. The method is based on the embedding model’s ability to capture the meaning of words from large collections of texts. The cosine measure is applied to calculate the distances between the word embedding’s location of the list of ethnic forenames and the indicators lists. The cosine is a distance measure applied more often in capturing distances between words in word embeddings and is therefore used to determine the association between words for this study (Kroon et al., 2021; Sorato et al., 2021; Wevers, 2019). Cosine distances are captured for all ethnic groups, resulting in vectors containing the distance scores between the ethnic group and the high-threat and low-status indicators. The cosine measure takes the cosine of the angle between two vectors. The closer the angle, the higher the cosine is and the closer the ethnic group is associated with the high-threat or low-status stereotype. The captured distance scores will be tested for significant differences among groups. Comparing these distances among ethnic groups can, if significant, indicate whether an ethnic group is perceived as having a lower status or perceived as being a higher threat compared to the other. This may signal a stereotype present in the corpus of the Dutch public news outlet’s articles. The formulas describe the cosine between two vectors; u and v as well as the measure of association s , with an ethnic name w and attribute list A . The indicators within the high-threat and low-status attribute lists are denoted with a . The over-right arrow equals the trained embedding model’s vector for the specific token.

$$COS(u, v) = \frac{u * v}{\|u\| \|v\|}$$

$$s(w, A) = mean_{a \in A} COS(\vec{w}, \vec{a})$$

A shapiro-Wilk test will be executed to test if the data is normally distributed. The result will indicate whether a parametric or non-parametric test should be applied. The Shapiro-Wilk test outputs a Shapiro statistic in combination with a p value. The null hypothesis states that the data is normally distributed. Thus, a p value less than α (.05) indicates that the null hypothesis is rejected, and the data does not come from a normal distribution.

4.2 Sentiment Analysis

The second research question; *How do sentiment distribution differ within articles that contain non-western ethnic names compared to articles that contain*

native and western ethnic names?, is investigated with the NRC lexicon’s sentiment score outputs and by comparing the distribution of articles that contain native names, western ingroup names and non-western outgroup names. Inspecting the sentiment distribution for differences will indicate how different ethnic groups are discussed in comparison to the other. For the sentiment analysis, the NRC lexicon is applied (Mohammad & Turney, 2013). The lexicon contains a list of words with associations connected to the 8 basic emotions of Plutchik (1980); joy, sadness, anger, fear, disgust, trust, anticipation, and surprise.

The following formula describes the calculation of the emotion scores per article. $Score(d, e)$ denotes the emotion scores e for article d . The denoted $sLexicon(w, e)$ refers to the emotion, association e for word w in the lexicon. The emotion scores per article equals the sum of emotions of the words w , in article d that exist in the lexicon.

$$Score(d, e) = \sum_{w \in d} sLexicon(w, e)$$

The sentiment analysis is performed on a collection of articles that include the ethnic forenames from the ethnic name lists. The articles are then ordered by country of origin and each article’s NRC scores are normalized. To see if there are differences in article sentiments between articles that discuss different ethnic groups, the polarity scores and emotion scores of all articles per ethnic group will be compared. The significance of the findings will be determined with statistical testing. If there are significant differences in polarity and emotion scores, it will indicate how different ethnic groups are discussed in comparison to the other.

A shapiro-Wilk test will be executed to test if the data is normally distributed. The result will indicate whether a parametric or non-parametric test should be applied.

5 Results

5.1 Stereotype Proximities

To investigate stereotype differences between the native group, ingroup- and outgroup members, a single word embedding model was trained on the entire corpus of NOS articles. The descriptives of the captured cosine distances between ethnic groups and high-threat and low-status stereotype dimensions are shown in table 2, table 3 and figure 1. Table 2 shows the ingroup members (*Greece* and *Italy*) to be relatively more neutral in comparison with the outgroup members (*Iran*, *Morocco*, *Turkey*, *Syria*, *Somalia* and *Afghanistan*) and the native group (*the Netherlands*) within the low-status dimension. Syria seems to be the exception here having a relatively high distance with low-status indicators in comparison with the other outgroup members. Table 3 shows that Greece’s distance is the smallest of all ethnic groups, indicating a close proximity to the high-threat association. This stands out as Greece is an ingroup member and is therefore expected to be relatively more on the neutral side within the

stereotype dimensions. Furthermore, it stands out that the outgroup’s captured distances are smaller for both high-threat and low-status stereotype dimensions in comparison with the ingroup. However, it should be examined whether these differences are significant. Figure 1 visually presents the mean captured distances for both low-status and high-threat distances. The y-axis displays the countries’ and group’s distances to the low-status dimension. The x-axis displays countries’ and group’s distances to the high-threat dimension.

Table 2: Descriptives low-status cosine distances per country & group-membership.

	Mean	Median	Standard Dev.
Netherlands (native)	0.381	0.423	0.265
Italy (ingroup)	0.371	0.416	0.267
Greece (ingroup)	0.333	0.365	0.270
Iran (outgroup)	0.257	0.266	0.279
Morocco (outgroup)	0.307	0.355	0.283
Turkey (outgroup)	0.356	0.385	0.258
Syria (outgroup)	0.423	0.461	0.253
Somalia (outgroup)	0.298	0.340	0.288
Afghanistan (outgroup)	0.255	0.26	0.277
Outgroup	0.317	0.355	0.280
Ingroup	0.359	0.399	0.268

Table 3: Descriptives of high-threat cosine distances per country & group-membership.

	Mean	Median	Standard Dev.
Netherlands (native)	0.427	0.461	0.251
Italy (ingroup)	0.465	0.513	0.26
Greece (ingroup)	0.401	0.442	0.268
Iran (outgroup)	0.432	0.456	0.244
Morocco (outgroup)	0.43	0.491	0.278
Turkey (outgroup)	0.47	0.508	0.24
Syria (outgroup)	0.444	0.494	0.276
Somalia (outgroup)	0.448	0.483	0.251
Afghanistan (outgroup)	0.441	0.479	0.263
Outgroup	0.442	0.485	0.262
Ingroup	0.446	0.491	0.264

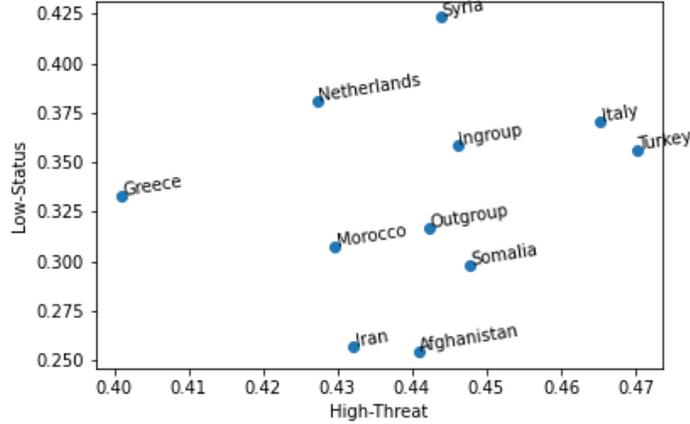


Figure 1: High-threat and low-status associations across ethnic groups.

A Shapiro-Wilk test is conducted to test for normality. Table 4 and table 5 shows its results. Table 4 shows a significant result for the three groups; native, outgroup and ingroup. This indicates the low-status vectors are not normally distributed. Table 5 shows a similar result indicating the high-threat vectors are also not normally distributed. Hence, a non-parametric significance test should be applied. The Kruskal-Wallis test, a test for multiple comparisons, is used to determine whether there are differences in measurements between ethnic groups. The Kruskal-Wallis is selected because it allows for testing for differences between 3 or more groups. When the Kruskal-Wallis outputs a significant p value, a follow-up test should be applied to test for differences across all of the included pairs to detect which of the pairs differ. For the ethnic groups' proximities to low-status indicators, the Kruskal-Wallis test indicates differences exist among the ethnic groups, $H(2) = 1178.51$, $p < .001$. Furthermore, the Kruskal-Wallis test indicates that the proximities with high-threat indicators also differ over ethnic groups as well, $H(2) = 81.11$, $p < .001$.

Table 4: Shapiro-Wilk test for normality on low-status distance vectors for different ethnic groups. p values; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

	Shapiro Statistic	p value
Native	0.954	***
Outgroups	0.977	***
Ingroups	0.970	***

Table 5: Shapiro-Wilk test for normality on high-threat distance vectors for different ethnic groups. p values; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

	Shapiro Statistic	p value
Native	0.971	***
Outgroups	0.959	***
Ingroups	0.958	***

A Dunn’s test is followed to identify what pairs’ differences is significant; native-outgroup, native-ingroup or outgroup-ingroup. A summary of the p values can be found in table 6. The p values are adjusted for multiple comparisons with use of the Bonferroni method. The method multiplies the original p value by the number of comparisons. The results show significant differences between all pairs; native-outgroup, native-ingroup and outgroup-ingroup. This result indicates all group pairs are different from another in their relation with low-Status stereotypes.

In addition, a Dunn’s test is also performed for the ethnic groups’ relation to high-threat indicators. The results of the Dunn test is summarized in table 7. Results are again corrected with the Bonferroni method. The results show significant differences between native-outgroup and native-ingroup, but not for outgroup-ingroup. This indicates the native group is significantly different in its relation with high-threat indicators in comparison with outgroups and ingroups. Furthermore, outgroups and ingroups seem to be not significantly different from each other within the high-threat stereotype dimension.

To conclude the first research question: *How do stereotypes differ in non-western ethnicities compared to western and native ethnicities?*. The hypotheses state the Dutch public news outlet to implicitly associate non-western outgroups closer with low-status indicators (H1) as well as with high-threat indicators (H2). The result of the Kruskal-Wallis and follow-up Dunn test for low-status indicators can be found in table 6. The results show differences exist between all paired comparisons: native-outgroup, native-ingroup and outgroup-ingroup. Figure 1 displays the native group to be the least associated with low-status stereotypes amongst the groups, followed by ingroups and outgroups respectively. This finding is in line with H1. Furthermore, table 7 shows the results of the Kruskal-Wallis and follow-up Dunn test for high-threat indicators. The results show differences exist between native and outgroups, native and ingroups but not between outgroups and ingroups. The examination of figure 1 reveals the native group to be the closest in proximity to the high-threat indicators. It should be noted that this finding contradicts the second hypothesis.

Table 6: Dunn’s post-hoc pairwise comparison, p values for ethnic groups and low-status stereotypes. p values; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

	adj. p value
Native - Outgroup	***
Native - Ingroup	***
Outgroup - Ingroup	***

Table 7: Dunn’s post-hoc pairwise comparison, p values for ethnic groups and high-threat stereotypes. p values; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

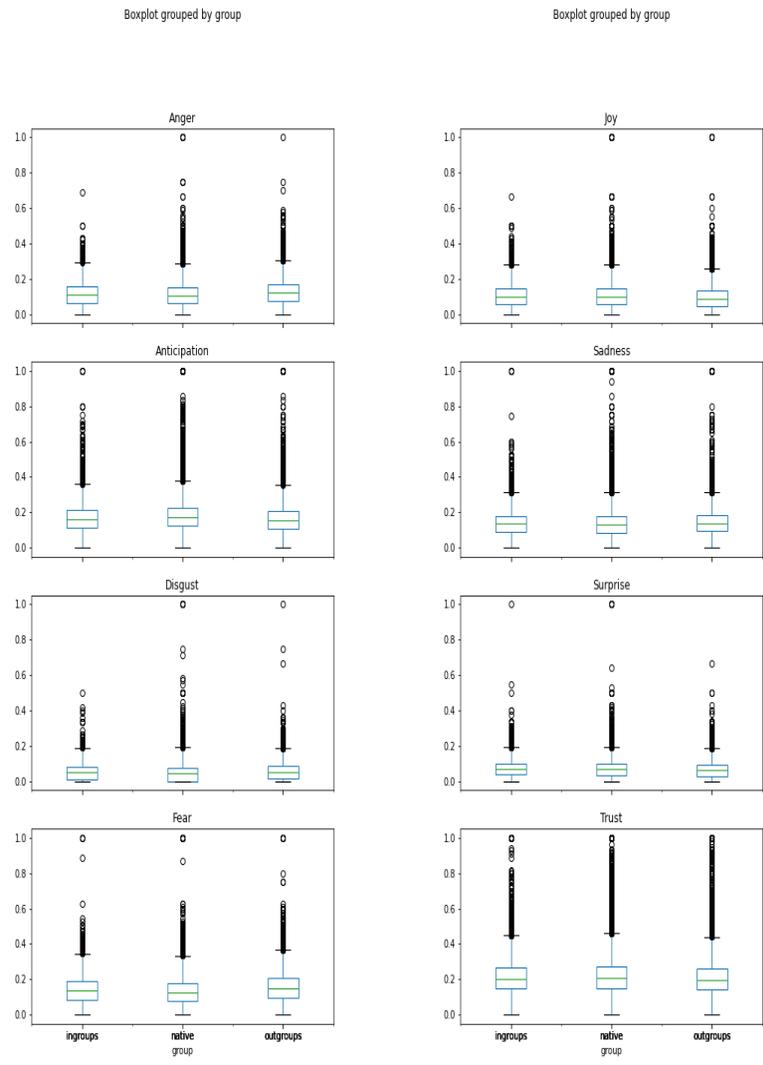
	adj. p value
Native - Outgroup	***
Native - Ingroup	***
Outgroup - Ingroup	0.08

For the first hypothesis, the result of the analysis indicate native groups to be closer in proximity to high threat indicators in comparison with out- and ingroups. Thus, evidence is found for hypothesis 1 where native groups are suggested to be implicitly less associated with low-status stereotypes in comparison with outgroups. Ingroups are also found to be implicitly less associated with low-status stereotypes. For the second hypothesis, the result of the analysis indicate native groups to be implicitly more associated with high threat indicators in comparison with out- and ingroups. Out- and ingroups seem to be equally associated with high-threat indicators. This finding in contrast with hypothesis 2 that stated outgroups to be implicitly more associated with high-threat indicators in comparison with native- and ingroups.

5.2 Sentiment Investigation

To investigate the sentiments displayed in the discussion of different ethnic groups, the National Research Council Canada (NRC) emotion scores are analyzed as well as the polarity score. Figure 2 shows the medians and the quartiles of the ethnic groups for all emotion scores. The figure shows that for the negative emotions; anger, disgust, fear and sadness, the native group has the lowest medians (.078, .03, .09 and .097). Herein, the ingroups follow second (.081, .031, .095 and .1) and the outgroup returns the highest medians (.087, .033, .103 and .101). Within the other emotions, it stands out that the native group has the highest median scores for anticipation and joy (.1 and .059). Herein, ingroups follow second (.095 and .057) and outgroups end at last (.091 and .052). For the remaining emotions; surprise and trust, ingroups score the highest (.041 and .12) followed by native (.04 and .12) and outgroups (.038 and .118). Figure 3 displays the median and the quartiles of the ethnic groups for the polarity measure. Inspecting the figure shows the median is the highest for the native group (.597) followed by ingroups (.579) and outgroups (.56), respectively. Sta-

tistical significance testing should reveal whether these differences are in fact significant.



19
Figure 2: A boxplot display of the sentiment distributions among groups.

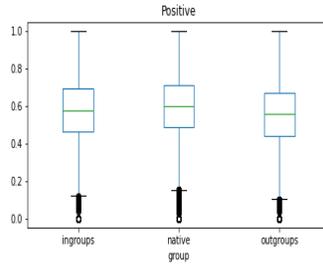


Figure 3: A boxplot display of the polarity distributions among groups.

For every emotion category and ethnic group, a Shapiro-Wilk test is performed to test for normality of the data. The resulting statistics and p values are presented in table 8. The table shows the emotion scores for every group to be not normally distributed. As a result, the Kruskal-Wallis, a non-parametric test to test for differences among groups, is applied. The Kruskal-Wallis test allows for testing for differences between 3 or more groups. The result of the Kruskal-Wallis test has indicated for every emotion category that significant differences exist among groups. To identify among which pair(s) the differences exists, Dunn's test for multiple comparisons is applied. To adjust for multiple comparisons, the Bonferroni method is selected.

Table 8: Shapiro-Wilk test for normality on NRC emotion scores for different ethnic groups. p values; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

Emotion	Group	Shapiro Statistic	p value
Anger	Native	0.955	***
	Outgroups	0.97	***
	Ingroups	0.971	***
Anticipation	Native	0.923	***
	Outgroups	0.941	**
	Ingroups	0.934	***
Disgust	Native	0.867	***
	Outgroups	0.905	***
	Ingroups	0.908	***
Fear	Native	0.965	***
	Outgroups	0.973	***
	Ingroups	0.956	***
Joy	Native	0.945	***
	Outgroups	0.935	***
	Ingroups	0.951	***
Positive	Native	0.993	***
	Outgroups	0.995	***
	Ingroups	0.993	***
Sadness	Native	0.919	***
	Outgroups	0.935	***
	Ingroups	0.938	***
Surprise	Native	0.926	***
	Outgroups	0.935	***
	Ingroups	0.911	***
Trust	Native	0.903	***
	Outgroups	0.896	***
	Ingroups	0.908	***

Table 9 displays the results of the Kruskal-Wallis and post-hoc Dunn test. For every comparison, the difference in emotion score is found to be significant. This indicates that for the found differences between the negative emotion; anger, disgust, fear and sadness, articles that discuss native groups show significantly less words that are associated with these emotions in comparison with ingroups- and outgroups. The outgroup is discussed with significantly the most negative emotions associated words. The opposite is found for the polarity score 'positive', the articles that discuss the native group members have significantly the highest positive score, followed by ingroups and outgroups, respectively. For the emotions anticipation and joy, it is found that the more frequent use of words associated with these emotions. is significant for native groups in comparison with outgroups and ingroups. In addition, articles that discuss ingroups are using more words that are associated with these emotions compared to articles that discuss outgroups. Finally, words indicating surprise and trust are signif-

icantly used the most in articles that discuss ingroups followed by native and outgroups, respectively.

Table 9: The results of the Kruskal-Wallis test and post-hoc Dunn test for multiple comparisons. Dunn’s p values; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

Emotion	Kruskal-Wallis H	p values	Native-Outgroup	Native-Ingroup	Outgroup-Ingroup
Anger	955.606	0.0	***	***	***
Anticipation	1200.544	0.0	***	***	***
Disgust	263.297	0.0	***	***	**
Fear	1350.954	0.0	***	***	***
Joy	716.671	0.0	***	***	***
Sadness	129.518	0.0	***	***	*
Surprise	259.226	0.0	***	*	***
Trust	123.868	0.0	***	***	**
Positive	1008.562	0.0	***	***	***

The hypothesis for the sentiment analysis stated that articles discussing outgroups have a more negative tone, more negative emotions and less positive emotion displayed in comparison with articles that discuss native or ingroups. The findings of this study are in line with this hypothesis. The words signalling the negative emotions are used more frequent within articles that discuss the outgroups compared to articles that discuss native and ingroups. Next to this, the polarity score indicates the same; More overall positivity is found within articles that contain native and ingroups compared to when outgroups are discussed.

5.3 Discussion on Results

The word embedding model has revealed ethnic outgroups to be closer in proximity to low-status indicators. Low-status indicators are words that are associated with (un)intelligence, low education levels, unemployment, addiction, and/or homelessness. A Kruskal-Wallis test indicated differences between at least two ethnic groups, Dunn’s post-hoc pairwise comparison confirmed differences exist between all pairs’ measurements. This investigation found outgroups to have the closest association with low-status indicators followed by ingroups and the native group, respectively. Furthermore, the word embedding model revealed ethnic outgroups not to be closer in proximity to high-threat indicators. High-threat indicators are words that are associated with hostility, deviance, threatening behavior or objects, criminality, judicial authorities, and/or illegal activities. A kruskal-Wallis test indicated differences between at least two ethnic groups, Dunn’s post-hoc pairwise comparison confirmed differences exist between the native group and outgroups, the native group and ingroups, but not between outgroups and ingroups. This investigation found the native group to have the closest association with high-threat indicators, outgroups and ingroups seem to have similar association with high threat indicators as no significance was found

in the pair.

To conclude the first research question: How do stereotypes differ between non-western, western and native ethnic groups. The results of this research indicate ethnic outgroups to be more associated with low-status stereotypes compared to native and ingroups and is thereby in line with the first hypothesis. This association is however not found for high-threat stereotypes. The result of the trained embedding analysis only indicate a partly ethnic bias towards outgroups as the closer association is found for low-status indicators that signal (un)intelligence, low education levels, unemployment, addiction, and/or homelessness. However, the closer association is not found for high-threat indicators associated with hostility, deviance, threatening behavior or objects, criminality, judicial authorities, and/or illegal activities. It stands out that for the high-threat stereotype dimension, the native group is found to be the most associated.

The sentiment analysis investigated different sentiment distributions between articles that contain native, ingroup and outgroup references. A Shapiro-Wilk test indicated none of the emotion scores to be normally distributed. Hence, the non-parametric Kruskal-Wallis test is applied on all emotion measures to investigate if significant differences exist between ethnic groups. Dunn's post-hoc pairwise comparisons confirmed differences exist within every emotion distribution as well as for the polarity score. All negative emotions; anger, disgust, fear and sadness are found to be more present in articles discussing outgroups followed by ingroups and the native group respectively. The positive emotion joy is found to be the least present in articles that discuss outgroups and more present in articles that discuss ingroups and the native group, respectively. Furthermore, the polarity score positive is found to be the lowest in articles that discuss outgroups, followed by ingroups and the native group, respectively. All results of the sentiment analysis confirm the third hypothesis that stated articles that contain non-western ethnic names have a more negative tone, include less expressed positive emotion and more expressed negative emotion than articles that contain native- and western ethnicity names. The results indicate articles that contain outgroup members are more negatively toned, contain more negative expressed emotion and less positive expressed emotion in comparison with ingroups and the native group. The results of the second research question thereby indicate an ethnic bias present within the articles of the Dutch public news outlet.

The results of the first research question are partly in line with the studies of Kroon et al. (2021) and Sorato et al. (2021) who both found ethnic outgroups to be closer associated with the stereotype indicators high-threat and low-status/drug use, prostitution, crimes, and poverty, and thereby found ethnic bias might be present within the investigated Dutch and Spanish newspapers' articles.

The results of the second research question are not in line with the investigation of Guo and Zhang (2020) who found no difference in sentiment orientation between ethnic outgroups and ingroups. Also the articles of the Dutch public news outlet are found to have an overall positive tone, while the research of Guo and Zhang (2020) found articles were overall skewed to the negative. The results

do align the findings of Bleich and van der Veen (2021) who found articles that discuss Muslim groups to contain more negative tone compared to articles that describe Hindus, Jews and Catholics.

6 Conclusion

The aim of this study was to investigate ethnic bias within the NOS, the Dutch public news outlet. The applied methods include training a word embedding model on the news outlets' articles and investigating sentiments in the discussion of different ethnic group members, and applying sentiment analysis on articles that discuss ethnic groups. Ethnic ingroups, outgroups and the native group have been compared for differences in the embedding model's distances between the groups and stereotype dimensions. Furthermore, the ethnic groups are compared in their sentiment distributions to discover differences their portrayal within the news outlet's articles.

The results of this study show ethnic stereotype might be present within the Dutch public news outlet's coverage. The literature stated media portrayals can have a large impact on shaping or changing the opinion of people as well as the attitude one may have towards another (Garrett et al., 2016; Meirick, 2013). Hence, the importance of fair representations of minority groups within media. The results of the first research question signal a closer implicit association with low-status stereotypes for outgroups in comparison with native and ingroups. The results of the second research question signals more negative, less positive and an overall more negative tone to be present in articles that discuss outgroup members. The stereotypical association in combination with implicit negative portrayal may signal ethnic bias and can feed negative prejudice against outgroups. The newspapers' portrayal of ethnic outgroups may therefore be subject to reconsideration to harvest more favourable attitudes towards them and create more inclusion in society.

7 Limitations and Future Work

This study has found evidence that may signal ethnic bias present in the Dutch public news outlet's coverage. The method that has led to the findings include the creation of a word embedding model from the corpus of pre-processed texts as well as the application of the NRC emotion lexicon. The findings are based on the complete collection of NOS articles from the period January 1st 2010 until May 1st 2022. The research did not focus on how and why the bias occurred. Therefore it remains unclear what topics have led to the biases found in this research. For example, a large and long covering of Dutch criminal trials may have created a bigger association with the native group and high-threat indicator list, while other trials gain less attention. However, this remains unclear because the reasons that lead to the result have not been investigated in this study.

Furthermore, the employed NRC emotion lexicon for sentiment analysis

comes with its limitations. Lexicon-based approaches are rule-based systems that base scores on word appearances in texts. The method struggles with negation in texts and can capture opposite scores from reality. In addition, the applied lexicon was translated from the English version. The translations in Dutch appear to be less rich than the English original. This is noted as Dutch words appear multiple times in the translated lexicon. For example, the Dutch translation of the English words 'appalling', 'frightful', 'horrible', 'horribly', 'terrible' and 'terrific' all equal 'verschrikkelijk'. The Dutch version of the NRC emotion lexicon may be less adequate in detecting different emotional words. In addition, the performed sentiment analysis score was applied to whole articles. As a result, a possible extreme display of emotions in parts of the article might have failed to detect. Lastly, displayed negativity or positivity can also exist within a combination of words. The method was only constructed to capture sentiment related to single word uses. Future research could consider methods that also detect specific parts within articles, can derive sentiments from combinations of words and can deal with negations in texts.

Both methods come with their limitations and these must be taken into account when examining the results in this study. However, this paper and its findings contribute to the literature on the computational identification of ethnic bias within newspaper articles and its findings may motivate further investigation to the cause of the found indicated ethnic biases.

The results of this study are reproducible as the notebooks containing the calculations and the used files are available on GitHub ⁶. At last, with regards to the ethical considerations, it is to the best of our knowledge that there were no ethical violations regarding the use of the dataset. All articles that are analyzed are publicly available from the NOS website ⁷ and are therefore believed to be available for anyone who wishes to analyze its contents.

⁶https://github.com/carpedimi/Ethnic_Bias_Dutch_Public_News

⁷<https://nos.nl>

References

- Arendt, F., & Karadas, N. (2017). Content analysis of mediated associations: An automated text-analytic approach. *Communication Methods and Measures*, *11*(2), 105–120.
- Arendt, F., & Northup, T. (2015). Effects of long-term exposure to news stereotypes on implicit and explicit attitudes. *International Journal of Communication*, *9*, 21.
- Bharathi, S. S., & Geetha, A. (2019). Determination of news biasedness using content sentiment analysis algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*, *16*(2), 882–889.
- Bleich, E., & van der Veen, A. M. (2021). Media portrayals of muslims: a comparative sentiment analysis of american newspapers, 1996–2015. *Politics, Groups, and Identities*, *9*(1), 20–39.
- 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. *Advances in neural information processing systems*, *29*.
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186.
- Chaabani, Y., Toujani, R., & Akaichi, J. (2018). Sentiment analysis method for tracking touristic reviews in social media network. In *International conference on intelligent interactive multimedia systems and services* (pp. 299–310).
- Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of machine learning research*, *12*(ARTICLE), 2493–2537.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2018). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. In *Social cognition* (pp. 162–214). Routledge.
- Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16), E3635–E3644.
- Garrett, R. K., Weeks, B. E., & Neo, R. L. (2016). Driving a wedge between evidence and beliefs: How online ideological news exposure promotes political misperceptions. *Journal of Computer-Mediated Communication*, *21*(5), 331–348.
- Gorodzeisky, A. (2013). Mechanisms of exclusion: Attitudes toward allocation of social rights to out-group population. *Ethnic and Racial Studies*, *36*(5), 795–817.
- Gorodzeisky, A., & Semyonov, M. (2016). Not only competitive threat but also racial prejudice: Sources of anti-immigrant attitudes in european societies. *International Journal of Public Opinion Research*, *28*(3), 331–354.
- Guo, S., & Zhang, G. (2020). Using machine learning for analyzing sentiment

- orientations toward eight countries. *Sage Open*, 10(3), 2158244020951268.
- Jacobs, L., Damstra, A., Boukes, M., & De Swert, K. (2018). Back to reality: The complex relationship between patterns in immigration news coverage and real-world developments in dutch and flemish newspapers (1999–2015). *Mass Communication and Society*, 21(4), 473–497.
- Kentmen-Cin, C., & Erisen, C. (2017). Anti-immigration attitudes and the opposition to european integration: A critical assessment. *European Union Politics*, 18(1), 3–25.
- Kroon, A. C., Trilling, D., & Raats, T. (2021). Guilty by association: Using word embeddings to measure ethnic stereotypes in news coverage. *Journalism & Mass Communication Quarterly*, 98(2), 451–477.
- Kroon, A. C., Van Selm, M., Ter Hoeven, C. L., & Vliegthart, R. (2018). Reliable and unproductive? stereotypes of older employees in corporate and news media. *Ageing & Society*, 38(1), 166–191.
- Meirick, P. C. (2013). Motivated misperception? party, education, partisan news, and belief in “death panels”. *Journalism & Mass Communication Quarterly*, 90(1), 39–57.
- Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436–465.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion* (pp. 3–33). Elsevier.
- Rekabsaz, N., West, R., Henderson, J., & Hanbury, A. (2021). Measuring societal biases from text corpora with smoothed first-order co-occurrence. In *Proceedings of the international aaai conference on web and social media* (Vol. 15, pp. 549–560).
- Ren, G., & Hong, T. (2017). Investigating online destination images using a topic-based sentiment analysis approach. *Sustainability*, 9(10), 1765.
- Ruigrok, N., & Van Atteveltdt, W. (2007). Global angling with a local angle: How us, british, and dutch newspapers frame global and local terrorist attacks. *Harvard International Journal of Press/Politics*, 12(1), 68–90.
- Small, M. L., & Pager, D. (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives*, 34(2), 49–67.
- Sorato, D., Zavala-Rojas, D., & Ventura, M. d. C. C. (2021). Using word embeddings to quantify ethnic stereotypes in 12 years of spanish news. In *Proceedings of the the 19th annual workshop of the australasian language technology association* (pp. 34–46).
- Starosta, K., Budz, S., & Krutwig, M. (2019). The impact of german-speaking online media on tourist arrivals in popular tourist destinations for europeans. *Applied Economics*, 51(14), 1558–1573.
- 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.
- Van Atteveltdt, W. H. (2008). Semantic network analysis: Techniques for extracting, representing, and querying media content.
- Van Klingeren, M., Boomgaarden, H. G., Vliegthart, R., & De Vreese, C. H. (2015). Real world is not enough: The media as an additional source

- of negative attitudes toward immigration, comparing denmark and the netherlands. *European Sociological Review*, 31(3), 268–283.
- Wevers, M. (2019). Using word embeddings to examine gender bias in dutch newspapers, 1950-1990. *arXiv preprint arXiv:1907.08922*.
- Wiebe, J. M. (1994). Tracking point of view in narrative. *arXiv preprint cmp-lg/9407019*.
- Zschirnt, E., & Ruedin, D. (2016). Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies*, 42(7), 1115–1134.

Appendix

Dutch high-threat words and their english translation: afperser (blackmailer), agent (cop), agente (cop), arrestant (detainee), arrestanten (detainees), autodief (car jacker), autokraker* (car cracker), bajesklant* (jailer), bandiet (bandit), bandieten* (bandits), bankovervaller (bank robberer), bankrover (bank robberer), bedelaar* (beglar), bedreiger (threat), bende (gang), bendeleden (gang members), bendeleider (gang leader), bendelid (gang member), benden* (gangs), bendes (gangs), beroeps crimineel* (career criminal), berovingen (robberies), beschieting (shooting), beul (executioner), boef (crook), bolletjesslikker (drug swallower), bolletjesslikkers (drug swallower), bommenmaker (bomb maker), bordelen (brothels), brandstichter (arsonist), brandstichters (arsonists), corrupt (corrupt), criminaliteit (criminality), crimineel (criminal), criminelen (criminals), cyberpesten (cyberbullying), dader (offender), daders (offenders), delinquent* (delinquent), delinquenten (delinquents), dief (thief), draaideurcrimineel* (revolving-door criminal), drugsbaas (drug lord), drugsbaron (drug lord), drugsbende (drugs gang), drugsbendes (drugs gangs), drugs crimineel (drugs criminal), drugsdealer (drugs dealer), drugsdealers (drugs dealers), drugsgebruikers (drugs users), drugshandelaar (drugs dealer), drugshandelaars (drugs dealers), drugssmokkelaar (drugs smuggler), dubbelagent (double-agent), fietsendief* (bicycle thief), gangster* (gangster), gangsterbende* (gangster gang), gedetineerde (inmate), gedetineerden (inmates), gegijzelden (hostages), gevangbewaarders (prison guards), gevangene (prisoner), gevangenen (prisoners), gevangenisbewaarder* (prison guard), gevangnissen (prisons), geweldsman* (violent man), gijzelaar (hostage), gijzelaars (hostages), gijzelnemer (hostage taker), gijzelnemers (hostage takers), handlanger (accomplice), hardrijder* (speeder), hoofdagent (prime suspect), hoofdagente* (chief agent), hoofddader (main offender), hoofdverdachte (prime suspect), hooligan (hooligan), hooligans (hooligans), huurmoord* (contract murder), huurmoordenaar (contract killer), illegalen (illegal immigrants), inbreker (burglar), indringer (intruder), jeugdbende (youth gang), jeugdbendes (youth gangs), jeugddelinquent (juvenile delinquent), kaper (hijacker), kapers (hijackers), kidnapper (kidnapper), kidnappers (kidnappers), kinderlokker* (child molester), kindermisbruiker (child abuser), kindermoordenaar (child murderer), kindslaven (child slaves), krijgsgevangenen (prisoner of war), kruimeldief (petty thief), kunstdief* (art thief), ladykiller* (lady killer), lastpak (troublemaker), lastpost* (nuisance), liquidatie (liquidation), loverboy (loverboy), lovergirls* (lovergirls), lustmoordenaar* (lust killer), maffia (mafia), maffiabaas (mafia boss), maffiosi (mafia), maffioso (mafioso), maftabaas* (mob boss), massamoordenaar (mass murderer), massamoordenaars (mass murderers), mededader (accomplice), medegegetineerde* (fellow inmate), medegevangene (fellow inmate), medeplichtige (accomplice), medeverdachte (co-defendant), mensenhandelaren (human traffickers), mensensmokkelaar* (people smuggler), mensensmokkelaars (people smugglers), messentrekker* (knife puller), misdaden (crimes), misdadig (criminal), misdadiger (criminal), misdadigers (criminals), misdadigerwapenhandelaar* (thug arms dealer), moordenaar (killer), moordenaars (killers), moordenaar* (killer), moordenaars* (killers),

moordmachine (killing machine), moordverdachte (murder suspect), motoragent (motorcycle cop), neerstak (stabbed), neersteken (stab), ontvoeringen (kidnappings), oplichter (crook), overvaller (robber), pedofter* (pedophile), pedofteren* (pedophiles), piraten (pirates), plunderaar* (looter), plunderaars (looters), politieagent (police agent), politieagente (police agent), politieagenten (police agents), politiecommandant (police commander), politiegeneraal* (police general), politiegewonde* (injured police), politieman (police man), politiemannen (police men), politiemensen (police officers), politieofficier* (police officer), politiepost (police post), politierechercheur* (police detective), politiestaat (police state), politievrouw (police woman), poltiemensen* (police officers), pyromaan (pyromaniac), recidivist (repeat offender), relschopper* (rioter), relschoppers (rioters), roofmoord (robbery murder), roofoverval (robbery), scherpschutter (sniper), schutter (shooter), seriemoordenaar (serial killer), skinheads (skinheads), slaaf (slave), slachtoffers (victims), slaven (slaves), sluipschutter (sniper), sluipschutters (snipers), smokkelaar (smuggler), smokkelaars (smugglers), snelheidsduivel* (speed demon), souteneur* (pimp), stalker* (stalker), straatbende* (street gang), strafbaar (punishable), strafklacht* (criminal complaint), struikrover* (highwayman), tasjesdief* (purse thief), terreurgroep (terror organisation), terreurverdachte (terror suspect), terrorist (terrorist), uitbuiting (exploitation), vechtersbaas* (fighter), veelpleger (frequent offender), veelplegers (frequent offenders), veiligheidsagent (security guard), veiligheidsagenten (security guards), veiligheidspolitie (security police), verdachte (suspect), verkrachter (rapist), vermisten (missing persons), voortvluchtige (fugitive), vreemdeling (stranger), vrouwenhandelaar (human trafficker), wapenhandelaar (arms dealer), winkeldief (shop thief), winkeldievegge* (shop thief), wreker* (avenger), wurgmoord (strangulation), zakkenrollers (pick pockets), zedendelinquent (sex offender), zedendelinquent* (sex offender), zelfmoordenaar (suicide bomber), zwartrijder* (transport user who does not pay).

Dutch low-status words and their english translation: achterlijk (retarded), achterlijke* (retarded), achterstanden (being behind), achterstandskinderen* (disadvantaged children), achterstandsleerling* (disadvantaged pupil), achterstandsleerlingen (disadvantaged pupils), achterstandswijken* (deprived areas), achterstandwijken* (deprived areas), achterstelling* (deprivation), alcoholicus* (alcoholic), alcoholist* (alcoholic), alcoholiste* (alcoholic) , alcoholisten (alcoholists), analfabeet (illiterate), analfabete (illiterate), analfabeten (illiterates), armoedig (poor), barbaars (barbaric), bastaardzoon* (bastard son), bedelaar* (beggar), bedelaars (beggars), bijstandsgerechtigden* (welfare recipients), bijstandsgerechtigten (welfare recipients), boerenlul* (peasant), dakloze (homeless person), daklozen (homeless persons), dronkelap* (drunkard), druggebruikers* (drug users), drugsgebruiker (drugs user), drugsgebruikers (drugs users), drugsrunners* (drugs runners), drugstoeristen (drugs tourists), drugsverslaafde (drugs addict, drugsverslaafden (drugs addicts), hangjongere* (loiterer), hangjongeren (loiterers), hoer (whore), hoerenlopers* (whore runners), hulpbehoevend (in need of help), hulpbehoevende (in need of help), idiot* (idiot), junk (junk) , junkers* (junkies) , junks (junks), kansarme (under privileged) kansarmen* (un-

der privileged), kindertehuizen (children's homes), krottenwijk* (slum), laaggeschoold, (low-educated), laaggeschoolden (low-educated), laagopgeleide (low-educated), laagopgeleiden (low-educated), loser* (loser), malloot* (moron), minderwaardig* (inferior), nestbevuiler* (nest polluter), nietsnut* (good-for-nothing), onderklasse (underclass), onderontwikkeld (underdeveloped), ongeletterde* (illiterate), ongeschoolde (uneducated), overlastgevende (a nuisance), pooier (pimp), primitief (primitive), probleemjongeren (troubled youth), prostituee (prostitute), prostituees (prostitutes), prostitutiebedrijven* (prostitute companies), reljongeren* (riot youth), schoolverlaters (school drop-outs), slet* (slut), sloeber (slut), sloebers (sluts), spijbelaar* (truant), spijbelen (to skip class), straatarm (poverty-stricken), straatkinderen (street children), straatprostitutie* (street prostitute), sukkel (loser), taalachterstand (language delay), tienermoeder (teen mom), tienermoeders* (teen moms), uitkeringgerechtigden* (social benefit recipients), uitkeringsgerechtigden (social benefit recipients), uitwas* (excrescence), verschopelingen* (outcasts), verslaafde (addict), verslaafden (addicts), weeskinderen (orphan children), werkloos (unemployed), werkloze (unemployed), werklozen (unemployed), werkschuwe* (work shy), zwerver (vagrant), zwervers (vagrants)

Words denoted with an asterisk (*) appeared in the original indicator list but did not exist within the corpus' vocabulary.