Article

Similar citizen portrayals? Converging media-based othering in tabloids and broadsheets

Stefanie Walter [ORCID]
Technical University of Munich, Germany

Zoltán Fazekas
Copenhagen Business School, Denmark

Abstract
Political communication researchers studying the news media coverage often distinguish between broadsheets and tabloids when sampling relevant news outlets. But recent work has pointed towards a 'tabloidization' of news coverage, complicating the empirical distinction between the two. Computational methods for text analysis can help us better understand how distinct the news coverage between these two types of news outlets is. We take the Brexit referendum as a case study illustrating various aspects in which broadsheets and tabloids cover an issue permeated by othering and divisive rhetoric. We focus on Brexit-related news coverage before and after the referendum (N = 32,946) and use word embeddings to analyze the portrayal of different groups of citizens that can generate an in- and outgroup divide. First, we document the presence of media-based othering in the form of overly similar migrant and European Union citizen representations that are, in turn, very dissimilar to the UK citizen representation. Second, we show partial convergence between tabloid and broadsheet newspapers, as differences in the degree and characteristics of media coverage are rather small and specific.

Keywords
Word embeddings, text analysis, othering, citizens, tabloids, Brexit

Corresponding author:
Stefanie Walter, Department of Governance, Technical University of Munich, Richard-Wagner-Str. 1, Munich 80333, Germany.
Email: Stefanie.Walter@tum.de

Introduction

A few days before the United Kingdom's referendum on the withdrawal of its membership in the European Union (short: Brexit), the online edition of the *Daily Mail* published an article entitled "The true cost of our open borders revealed". It informed their readers about an "alarming" report showing that European Union (EU) migrants living in Britain are more likely to have jobs than British citizens. In addition, they are also more likely to be working in the United Kingdom (UK) compared to migrants from outside Europe. According to the newspaper, this finding supports the "leave" campaign, as it shows that the British "immigration system is 'racist' because it favors people from the continent over others" (Dathan, 2016).

This news story is noteworthy for several reasons. First, it contrasts and creates a dual divide between citizen groups: a) the ingroup of UK citizens and the outgroup of immigrants and b) within the outgroup a divide between EU and non-EU migrants. This is interesting, because both in- and outgroups are often conceptualized as being homogenous (cf., Share, 2018). While the news story overall does not seem to be explicitly negative in tone, the contrasting of the groups nevertheless conveys the message that UK (and non-EU) citizens are disadvantaged because of intra-EU migration. The aim of this paper is to study the news media's role in the exclusion of others by portraying them as systematically different from the ingroup.

To this end, we draw on the framework of 'othering' and utilize recent advances in computational methods for text analysis to examine news related to the 2016 Brexit referendum in the UK where these groups played an important role. We offer three distinct substantive contributions. First, we argue that 'othering' relies on similarities and differences between the portrayal of people: The exclusion of others is a function of the dissimilarity between in- and outgroups (heterogeneity). By comparing within group differences in the portrayal of EU citizens and non-EU migrants,[1] we examine whether the news media coverage reflects hierarchies or establishes homogeneity of outgroups.

Second, we assess between-media differences regarding the extent of 'othering' present in the media coverage and evaluate the role of news outlets in the spread of media-based othering and offer a comparison of the news coverage of tabloids and broadsheets. Our findings are especially relevant as the media coverage is a decisive factor for othering and related phenomena such as the rise of populist movements, but how the news media may contribute to such sentiments and how these are conveyed by the news media is less well understood. Finally, our study also advances our understanding of the representation of foreign nationals beyond the distinctive category of migrants.

Brexit and the construction of 'the other'

Brexit has been studied from different angles: Nationalism, British Exceptionalism, Euroscepticism, and right-wing populism. These approaches have in common that they foreground the concept of 'the people' from which – in one way or another – others are excluded. In other words, society is divided into different groups. People have the need to identify shared attributes that distinguish their group from others and it is this group

membership that leads them to differentiate between "us" (the ingroup a person belongs to) and "them" (the outgroup of others) (Hogg, 2006). To maintain a positive image of one's own group, individuals draw comparisons to other salient groups (Ortiz and Behm-Morawitz, 2015: 95).

The concept of 'othering' focuses on the process through which difference and sameness are established (powell and Menendia, 2016). Othering can be defined as a "set of dynamics, processes, and structures that engender marginality and persistent inequality across any of the full range of human differences based on group identities" (powell and Menendia, 2016:17). It puts emphasis on the *discursive practices* used to establish otherness and points our attention to how the news media constructs the "us" and "them" by emphasize differences and creating distance between the two (cf. Poole and Sandford, 2002).

The news media play a crucial role for intergroup comparisons (Harwood and Roy, 2005; Mastro, 2003; Ortiz and Behm-Morawitz, 2015). News media can activate social identities and lead to "media-based othering, that is, the development of a schema-based ingroup or outgroup bias in the perception and evaluation of social phenomena" (Krämer, 2014: 55). While work in this area – especially in the European context – has often focused on Muslims being constructed as the other (cf. Bleich et al., 2015; Foner, 2015), othering as a concept can be useful for the study of any group-based differences.

Focusing on the Brexit referendum and the relationship between salient in- and outgroups, the aim of this study is to examine the news media's contribution to 'othering' understood as "creating social distance" between groups (powell, 2017). Brexit is particularly interesting for the study of media-based othering, because it was prone to foster a divide between the British and multiple outgroups. Depending on the political standpoint, the referendum on the UK's membership in the EU provided a unique opportunity to highlight the differences or sameness between the British and the EU citizens.

EU citizens can be seen as an outgroup with a special relationship to the ingroup. Even though citizens from other member states are foreigners, they share(d) the EU membership with UK citizens and might be perceived as culturally and socio-economically close (cf. Hagendoorn, 1995). At the same time, EU citizens might pose a greater threat to the ingroup compared to citizen from countries outside the EU due to their rights to move and reside freely within the territory of the member states. Consequently, every EU citizen might be perceived as a potential EU migrant. Some have even argued that the UK represents a special case compared to other EU member states, as its citizens are more afraid of Europeans than citizens from outside the EU (Freeden, 2017).

Othering through contrasting citizen portrayals in the news media

Extant research on the news media portrayal of in- and outgroups has largely focused on the latter, while a comparison of the two is missing. The media representation of the ingroup has been examined through measuring how people-centered news stories are (Rooduijn, 2014). With regards to the outgroup, research has often focused on media attention to the issue of migration (e.g., Tong and Zou, 2019). There are several examples

documenting that immigrants have been portrayed in a negative light (see: Eberl et al., 2018 for an overview), and that negative depictions of social groups and their perception as a threat can result in attitudinal changes (Atwell Seate and Mastro, 2016) as well as in anti-immigrant vote choices (Blinder and Allen, 2016a; Dinas & van Spanje, 2011). Furthermore, migrants are also frequently portrayed as a political, economic, or physical threat to the well-being of the ingroup (Tong and Zou, 2019), although Blassnig et al. (2019) found that the media only rarely employ explicitly exclusionary rhetoric. Instead, *distancing* the ingroup from the outgroup or the *contrasting of groups*, for example via the media coverage, might be crucial for establishing (and maintaining) otherness and a group divide (Grove and Zwi, 2006).

Building on these considerations, we argue that the exclusion of others in the media coverage can be conceptualized as having two layers: First, the portrayal of the ingroup is important, because it functions as a "standard" against which others can be compared (Reinemann et al., 2016: 21) and deviations or distances from this standard would thus marginalize or dismiss other groups from being part of the ingroup. Accordingly, dissimilarity in how otherwise identical people (i.e., workers) from the ingroup (i.e., British) and the outgroups (German or Indian) are portrayed by the media highlights properties deemed important and can build an image with contrasting others.

Second, even along the same dimension of exclusion of others there can be multiple, but related, outgroups, identified by provenience or other features creating distance or sameness. As difference can be hierarchical (Innes, 2010: 559), we can potentially find different degrees of otherness. For the media coverage of EU citizens, there are two options. News media may emphasize shared attributes between British and EU citizens, establishing them as a distinct group, but one that is closer to the ingroup than non-EU migrants (by that we mean any mention of migrants without any specific link to the EU or its member states or explicit reference to their non-EU provenience). But the opposite might also be the case. We know that among migrants, especially those from EU member states are seen as taking advantage of the British welfare system and are being accused of taking job from the British (Tong and Zou, 2019: 448). In this case, the distance between British and EU citizens would be larger compared to non-EU migrants.

Alternatively, media-based othering may present outgroups in a more homogeneous manner to establish their sameness. Based on previous research, we know that the news media, especially tabloids, do not differentiate sufficiently between different groups of foreign nationals (Moore and Ramsay, 2017), depicting them as "one unified immigrant entity" (Share, 2018: 32). Even though the UK's exit from the EU should not impact migration levels in general, but intra-EU migration, the news coverage suggested that there is a link between migration levels and EU membership (cf. Tong and Zou, 2019). Establishing EU citizens and migrants as a homogenous outgroup may aid this framing of the EU as an entry gate for non-migrants that was emphasized in the news coverage surrounding the EU's "refugee crisis" coinciding with the Brexit referendum (Tong and Zou, 2019).

These layers of dissimilarity-similarity can only be present when we have at least three well identified or separable groups, and they do not necessarily go hand-in-hand: A second outgroup can be presented differently from the ingroup, while it does not

necessarily have to be depicted in the same way as the other outgroup. Overall, these facets allow for a rich set of potential comparisons and an in-depth understanding of media-based othering, but also offer a clear analytical framework for the comparison of different news media outlets.

Hypothesis: Othering in tabloid and broadsheet media

After our theoretical considerations, we ask: *What type of media outlets are more likely to contribute to media-based othering*? The answer to this question sheds some light on whether we are witnessing a convergence or divergence regarding the normalization of othering discourses in the news media. The main focus is on the distinction between tabloid and broadsheet media.

First, these outlet types have different relationships with the political establishment. Tabloid media are more likely to criticize the political establishment in their coverage and to sympathize with anti-establishment parties (Akkerman, 2011). They are therefore more likely to amplify populist voices (Mazzoleni, 2008; Rooduijn, 2014) who promote othering through rhetoric around ingroup/outgroup divides. Qualitative work points to the importance of language used by tabloids that contributes to a "categorization into binary divisions of the world" (Conboy, 2006: 16) into 'us' vs 'them'. Broadsheet media, on the other hand, are expected to either offer opposition or ignore populist movements because of their close alignment with mainstream political parties, which should lead to less media-based othering in their coverage.

Second, the audiences to which these media cater differ: Tabloids traditionally try to establish a close linkage with their readership by representing their views and interests (Conboy, 2006). For this kind of group identity processes, the comparison to other (out) groups is crucial (Ortiz and Behm-Morawitz, 2015: 95). To this end, the readership of tabloids (i.e., the ingroup) is often contrasted with an outgroup. In fact, both groups might even be depicted as polar opposites (Conboy, 2006: 32). Their readership is also frequently established as the community affected (Matthews and Brown, 2012), e.g., by EU citizens or non-EU migrants entering the British job market.

Tabloids, especially in the UK, are furthermore often Eurosceptic, while broadsheet papers are more supportive of a more extensive European integration (see De Vreese et al., 2006; Startin, 2015). Broadsheet coverage was found to be generally more Europeanized, reporting more about affairs and actors from the EU as well as other member states than tabloids (Kleinen-von Königslöw, 2012), and paying more attention to EU citizens (Walter, 2017).

Comparative research shows the UK exhibits a more polarized news coverage of the issue of migration compared to other European countries (Berry et al., 2015: 10). Negative reporting about the issue of immigration was especially prevalent in the tabloid press (Moore and Ramsay, 2017). However, a recent study by Walter (2019) that looked specifically at the media portrayal of EU citizens and migrants in Brexit news found no evidence for a more negative portrayal in British tabloids compared to broadsheets. As this might vary once we move beyond the sentiment of the news coverage, we notwithstanding hypothesize that:

(H_1) Distance from ingroup hypothesis: Tabloids portrayed both EU citizens and non-EU immigrants more dissimilar to UK citizens than broadsheets.

As we are unsure whether we should expect media-based othering to result in homogeneity or difference between outgroups, we ask: *Are EU citizens portrayed as more similar or different to non-EU immigrants and does the degree of sameness/difference vary between broadsheets and tabloids?*

Before continuing with our empirical analysis, we highlight some final considerations regarding the Brexit coverage case selection. We study a contested period where issues related to immigration were salient and the UK was pitted against another elite entity, the EU. Accordingly, we should treat this as a likely case to find evidence for othering in the news coverage compared to other periods. However, this is not an aspect we can directly address. We will return to this point to elaborate on differences surrounding the campaign period to further assess any sort of intentionality in the use of media-based othering. Finally, and most importantly, this is also a likely case for finding evidence for between-outlet differences, exactly for the same reason rooted in polarization between the two sides of the debate where, besides the diverging media practices, tabloids were often on opposing sides to broadsheets.

Data and analysis

Newspaper articles

We analyze Brexit related newspaper articles published by British broadsheets and tabloids in 2016. Our focus on Brexit is guided by the goal to assure all citizen groups studied here are a somewhat salient component of political news coverage and there is potential for polarizing content coverage, given the simple dichotomy of the referendum choice. Overall, these contribute to a well-defined scenario where between outlet differences are most likely to appear.

To reflect this aim, we used the Lexis Nexis and Factiva data bases to search in British daily (on- and offline) newspapers for articles published in 2016 that mention *Brexit*, *membership referendum*, *EU referendum*, or *European Union referendum* in their headline. For this study, the distinction between broadsheets and tabloids is crucial and we relied on previous research to identify the relevant outlets (Blinder and Allen, 2016b; Gabrielatos and Baker, 2008; Startin, 2015). The Brexit term search reveals good face validity, with a high concentration of articles around the referendum date and maximum numbers on the date (see Supplementary Information (SI)). Our data contains 32,946 articles published by a broad range of outlets (N = 24), also in terms of their political leaning, throughout 2016 with an equally split between broadsheet and tabloid newspapers.[2] A detailed data summary is available in Supplementary Information1.

Identifying citizen mentions

The next step is to identify citizen mentions in these texts, which was done by searching for a list of phrases. We apply a dictionary, or more precisely a thesaurus, where UK

citizens, EU citizens and immigrants, and non-EU Immigrants are described by various words or combination of words. We refer to non-EU immigrants as any case where (1) provenience or background is not mentioned or (2) there is an explicit non-EU mention included prior or post immigrant mention. We have three main goals: To use a small set of terms that are clearly linked to citizen or people mentions at the conceptual level and thus (1) minimize false positives, such as references to products, (2) rely on a comparable set of terms across the different citizen groups, and (3) devise a flexible search that can be further extended if needed. Figure 1 shows the components of this search.

When referring to immigrants, we search for versions of migrant, foreigner, immigrant, but also for those cases where citizens are mentioned without a specific provenience but being highlighted as foreign. In this combination, the second part of the search bigrams starting with foreign will be versions of citizen, people, national. Most importantly, these terms will form also the common core of our UK and EU citizen mentions. The difference will be in the first terms entering these search components, clearly highlighting a specific provenience. For example, a search for EU citizens in a broader sense returned a successful hit if we found some combination of what is listed in EU (1) and second terms in the figure (such as "Romanian* national" or "EU* citizens"). Upon inspection and reading of articles, we have carried out several updates and additional steps to assure the quality of the search results, detailed in Supplementary Information 2.[3]

We often see formulations where EU citizens were directly referred to as immigrants, such as "EU immigrants" or "Polish immigrants", which we group as EU citizens and immigrants in our analysis, since there is a clear provenience listed. Once all searches and relabeling was carried out, as our hypothesis considers potential differences between tabloids and broadsheets, we also split and replace the tokens allowing us to identify the group of citizens mentioned and the news outlet simultaneously, such as an immigrant mention being "tokenmigbroad" if it appears in a broadsheet, but "tokenmigtabloid" when it appears in a tabloid. This allows us to estimate the models on the full corpus, capitalizing on all words, but generating group specific word embeddings for the core tokens tagged as such. Consistent for both outlet types, the ingroup of UK citizens were the one that received most attention before and after the referendum, with tabloids using more people references in general (see Supplementary Information 2).

Estimation and validation

To understand the extent of media-based othering, we concentrate on how these groups were described in relation to each other. We do this by utilizing word vector models (word embeddings from now on), a computational approach extensively developed in natural language processing (see e.g., Mikolov et al., 2013; Pennington et al., 2014). In simple terms, words are represented in a multidimensional vector-space and their position in this space is a function of their meaning. Their meaning is given by their context, i.e., what sort of company a word keeps, usually given by neighboring words. It is important for our study that words with similar meanings are closer in the vector space, and thus overall, across multiple dimensions, they will be more similar in their vector positions to words with which they share meanings. However, it is not a necessary condition for similar
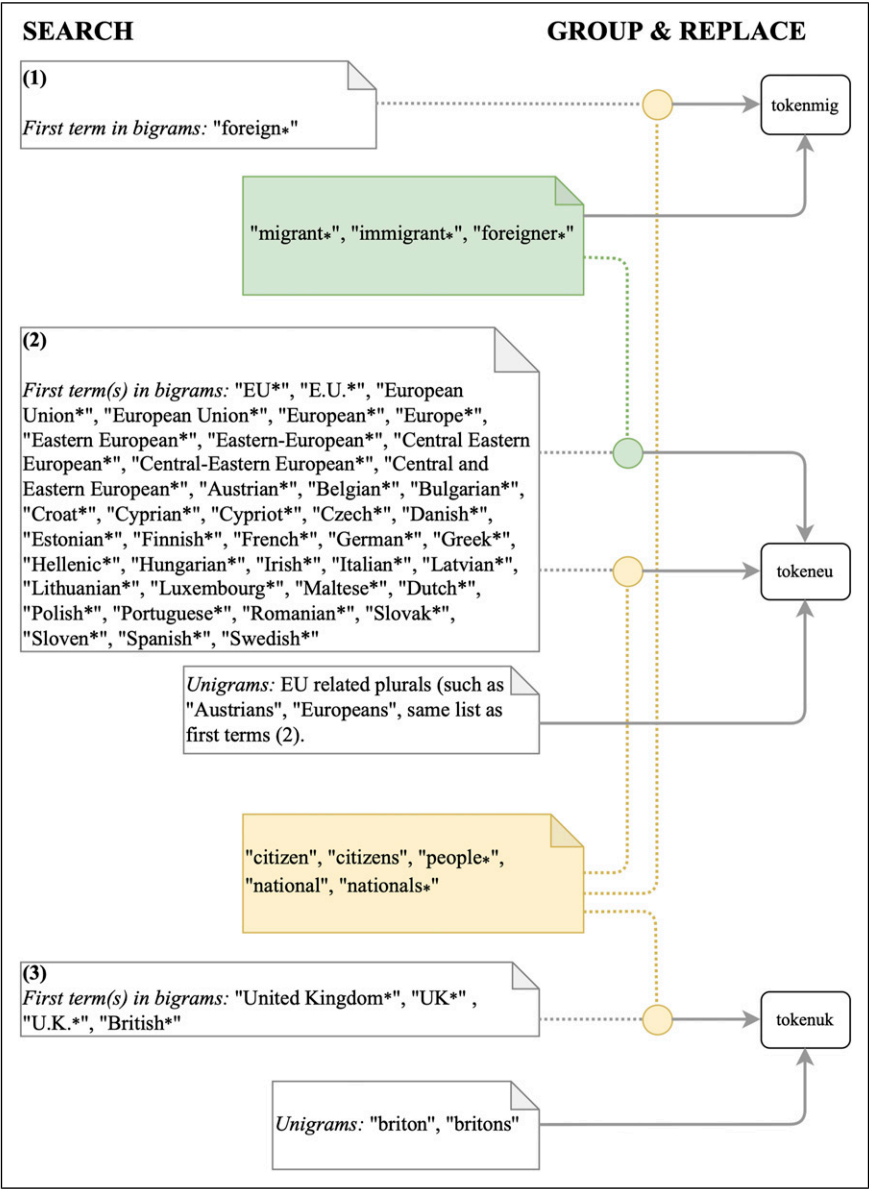
Figure 1. Content coding.

words co-occur at all: They might never show up together, but if they keep showing up in very similar contexts, they will still be similar in terms of meaning (see for detailed social sciences summary Kozlowski et al., 2019; Rodriguez and Spirling, 2021).

Overall, these approaches are extremely powerful to recover semantic relationships, thus they can also be used for detecting synonyms or finding syntactic variants of different words. These advantages also make them popular for social science and communication applications, being at the growing frontier of recovering social or political meaning or relationships using word embeddings, such as stereotype measurement (Garg et al., 2018), cultural categories (Nelson, 2021), or media and ethnic bias (Kroon et al., 2020). Finally, word embeddings are useful for comparisons between different corpora, or how the representations or meanings differ conditional on what texts we are analyzing, which allows us to contrast the media coverage in tabloids and broadsheets.

We employ a global log-bilinear regression model for the unsupervised learning of word embeddings (Pennington et al., 2014)[4], an implementation that further capitalizes on general word contexts. Since we are interested in the local context of the content categories, we use a window size of 6 words on both sides (with proximity weighting, see also Rodriguez and Spriling, 2021) and use 100 dimensions for the word vectors, with sensitivity tests carried out later. We retain the order of features within each document. Since Antoniak and Mimno (2018) showed that word embeddings and their similarity metrics can suffer from quite some instability especially with small corpora, we report results based on 50 bootstrapped samples, drawn with replacement from the corpus. This allows us to include uncertainty measures (confidence intervals) around the similarity estimates. We will follow a customary interpretation of these uncertainty bounds, however, it is worth noting that they are not related to any sort of representativeness ideas of our text corpus, rather they are a product of model fitting uncertainty.

We use the cosine-similarity between the resulting word vectors for the three groups of interest as the main quantity of interest. As a measure, it summarizes the size of the angle between vectors, ranging from 1 (vectors pointing in the same direction) to -1 (vectors pointing in the opposite direction), with around 0 values indicating independence. For preprocessing we converted the text to lower case and then removed punctuation, numbers, separators, and stopwords. We also removed terms that appear less than 5 times and carried out stemming, resulting in 30,255 unique stems for our final analysis.

Applications relying on word embeddings require additional validation (Bakarov, 2018; Rodriguez and Spriling, 2021). Our corpus is rather specific and thus standard evaluation tasks are not suitable since few terms from specific analogy tests are present in the corpus. Thus, we devise a custom analogy-based task that is corpus specific and we show in Supplementary Information 3 that our estimates have good face validity.

Results

Citizen portrayals in the media

In Figure 2, we summarize our main results by showing the cosine similarity between the estimated word vectors for the three citizen groups. As introduced, we rely on bootstrapping the corpus and generate multiple sets of word embeddings and on each of these sets we calculate the similarity scores. These are summarized using the line ranges (whiskers) in the plot, since we have a distribution of similarity scores based on the
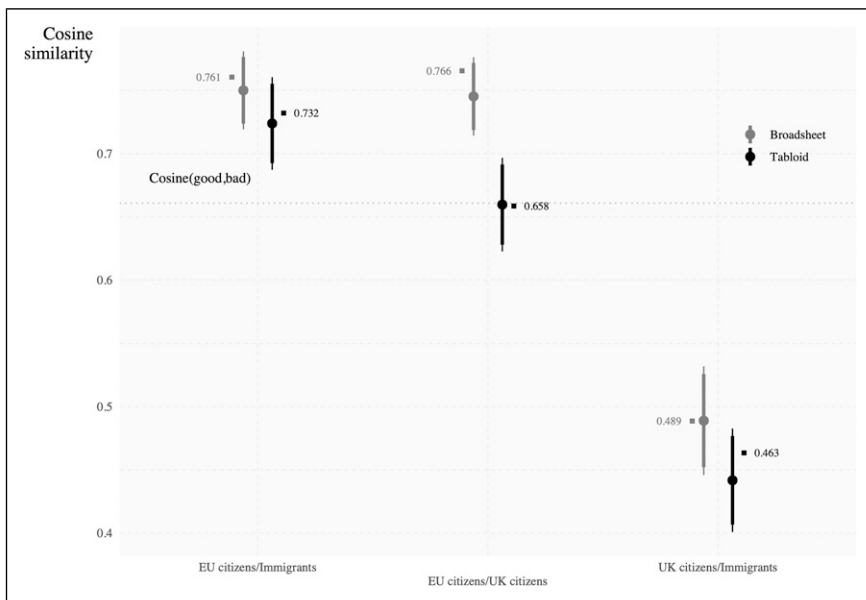
Figure 2. Cosine similarity of citizen mentions in broadsheets and tabloids.

multiple runs. We have also added the similarity estimates based on the original, full data (squares). The results are, naturally, the same when focusing on these numbers, however, our approach to add uncertainty assures that small differences are not overinterpreted. A useful heuristic in linking these visualizations to traditional significance interpretation would be to consider if the confidence bands of one estimate include another estimate's average.

The substantive clarification pertains regarding how we should actually interpret these similarity scores or differences between them, which is, undoubtedly, a difficult task. As we are working with a rather small and homogeneous corpus, which can create artificially high similarity scores, we aim to interpret these numbers exclusively in a comparative manner, between different groups or outlet types. We also display the cosine similarity between good and bad based on our full data, model, and estimation. This relatively high value underscores that we are not necessarily looking at valence related differences in these scores, rather higher similarity scores can be due to the use of these words in similar contexts. To be sure, some specific dimensions might reflect a valence related distinction (also picked up that on several of the 100 dimensions estimated they are quite far apart), but we are summarizing across all dimensions to have a broad comparison of terms and their use-based meanings. Thus, while good and bad are antonyms, they can be used to describe objects in a similar way, resulting in an often-shared context. For example, two different articles can talk about the economy saying that "the economy took a better/worse turn after the announcement of Brexit", and the choice of better or worse is very
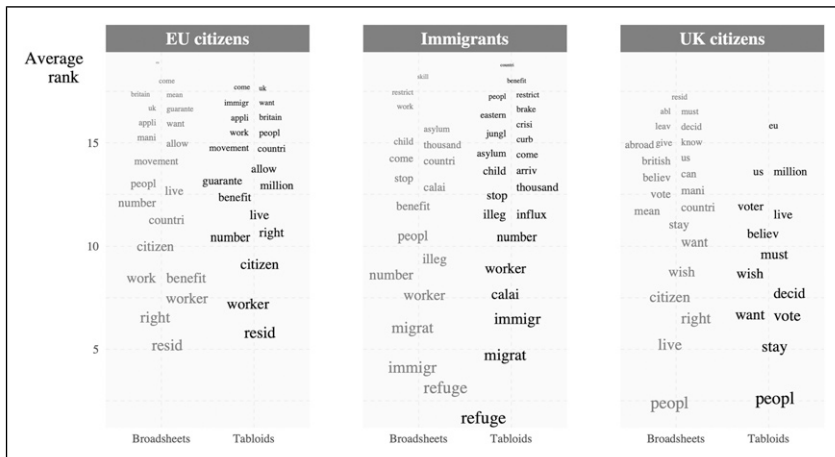
Figure 3. Most similar words compared.

important, but if used within these news articles in similar contexts, we will see overall larger similarity scores.

Figure 2 offers many insights, but first and foremost, we see patterns of othering through distancing the ingroup from the outgroup in the expected manner: Ingroup to outgroup similarity (UK citizens compared to EU citizens and UK citizens compared to [non-EU] immigrants) is lower than within-outgroup similarity (EU citizen compared to immigrants). This further informs us, substantively, how EU citizens were depicted in the Brexit coverage: More similar to immigrants, rather than to UK citizens. This phenomenon does not directly imply a negative or positive image, rather, it indicates that in the Brexit coverage EU citizens were depicted close to another customary outgroup, immigrants. However, the figures also show EU citizens being less of an outgroup, since they are more similar to UK citizens than are migrants to UK citizens.

The second main insight from our analysis summarized in Figure 2 is that the coverage of these different groups is quite similar in broadsheets (grey) and tabloids (black), with one notable exception. The between outlet-type differences are, on average, quite small: (1) the general ordering of similarity is identical and (2) EU citizen to migrant as well as UK citizen to migrant similarities are also quite close. One important exception is related to the EU citizen comparison with the ingroup. In broadsheets, EU citizens are depicted significantly more like UK citizens than in tabloids suggesting weaker media-based othering language expressed as distance between groups. In broadsheets, EU citizens are depicted as similar to UK citizens as they are to non-EU immigrants. Regarding the UK citizen and immigrant comparison, we see very similar patterns: Overall, immigrants are less similar to UK citizens in tabloids compared to broadsheet, while in both outlets, these two groups are the most dissimilar. Here, we note that these differences are substantively small. These findings confer partial support for our hypothesis expecting that tabloids

portrayed both EU citizens and immigrants more dissimilar to UK citizens than broadsheets.

In Figure 3, we offer a more qualitative picture of these results by looking at the most similar words to the citizen mentions. We see a rather strong overlap in the terms associated with these citizen groups between tabloids and broadsheets. Many of the EU citizen contexts are related to living and working rights and potential benefits that could come with that. Immigrants are also described through the numbers, movement and work, which is likely contributing to the strong overlap; however, we see here a stronger focus on similarities with refugees (and illegal status), which was rarely an important context or consideration when talking about EU citizens and EU immigrants. While the freedom of movement and working rights of UK citizens are likewise affected by the UK's exit from the EU, this seems to be to a larger extent reflected by the coverage of broadsheets, while tabloids seem to focus more on the referendum vote itself.

To reiterate, our findings can be interpreted as the way EU citizens have been talked about (including the topics and the descriptors) in comparison to UK citizens is more different than the same comparison of EU citizens and immigrants with (non-EU) immigrants. Talking about different groups in divergent contexts can fuel why we think of them (the outgroup) as being different, since we have fewer things in common with them. This manifestation can underscore an 'us' vs 'them' narrative by delimiting the groups. However, there is still similarity between EU and UK citizens, which can reflect common contexts and descriptors, which is not surprising, given economic and social topics have been relevant for both of these groups. But when presented and evaluated in a comparative manner, this overlap is still weaker than that with immigrants, and it is also weaker in tabloids compared to broadsheets.

To conclude, we have shown that, between-citizen mention similarities follow the expected broader patterns, indicative of media-based othering and the positioning of EU citizens as immigrants, rather than a closely comparable group with the ingroup. However, overall qualitative similarities are quite large, underlining the idea that all three groups are "people". We also find marked convergence between outlet types, with the exception of more distancing of EU citizens from the ingroup in the tabloid coverage of Brexit.[5]

Robustness checks

In the remaining part of our analysis section, we evaluate a set of alternative specifications to verify that the convergence between outlet types is not conditional on our data and methods related choices (see Supplementary Information 5). Keeping with the global log-bilinear model estimation, we altered the context window (6–3). Furthermore, journalists working for tabloids and broadsheets cover the same events and often rely on the same source of information, for example, a press release or a report by a news agency. Thus, we refitted our model on a subset of the corpus that excludes duplicates or very similar articles between broadsheets and tabloids. We further refitted our model where instead of the news article, we keep the sentence as our document level, to assure that we never include

context that is not within the same sentence of a word. Our results are essentially un-changed if we apply these transformations at the input data level or model parameters.

In addition, while using a global log-bilinear model to estimate the embeddings usually comes with more stability (Rodriguez and Spriling, 2021), we refitted our models using Word2Vec (Mikolov et al., 2013), relying on skip-grams, which can result in differences as the Word2Vec implementation is more prone to select rare terms, whereas these carry less weight the GloVe results. The differences between tabloid and broadsheet similarity scores are as follows: EU citizen/Immigrant −0.03 (with main result being 0.026); EU citizen/UK citizen 0.048 (main result 0.086), and Immigrant/UK citizen at −0.066 (main result 0.047). This is the only comparison where we find some divergence between the models, however, this is not necessarily surprising as other comparisons also reveal somewhat weaker correlations between the results from the two implementations (Rodriguez and Spriling, 2021).

First and foremost, the only comparisons from our main analysis that revealed systematic differences between tabloids and broadsheets (EU citizens vs UK citizens) are robust to the change of implementation approach. Second, the EU citizen vs immigrant difference, while opposite in sign, is still of very small magnitude. The only major difference appears at the UK citizen vs immigrant similarity scores, specifically driven by a low similarity based on the Word2vec model for broadsheets.

Substantively, the similarity or dissimilarity of the portrayal of in- and outgroups might vary depending on whether or not we are looking at a campaign period, and it could be that tabloids and broadsheets are influenced differently by the campaign coverage logic.

The second factor that could be related to the presence of othering might be the degree of political parallelism in a country, where parallelism is defined as the extent to which political advocacy is seen as part of the role of journalism and whether there are ties between newspapers and political parties (Brüggemann et al., 2014; Hallin and Mancini, 2004). During the Brexit referendum, national newspapers picked sides and gave explicit voting advice to their readers as to whether the UK should remain or leave the EU (e.g., Levy et al., 2016). Accordingly, the second updated differentiation concerns the media outlet endorsement choices and their potential overlap with the outlet type. Tabloids mostly advocated a leave position, whereas broadsheets supported the remain side, resulting in a strong overlap between outlet type and endorsement.

The results displayed in Figure 4 show the same general overall pattern as our main analysis did. There is a strong convergence between the outlets post-referendum, thus the larger differences regarding EU citizen and ingroup similarity are mostly present in the pre-referendum period. This indicates some evidence for the potential role the campaign had regarding the activation of othering within tabloids. When we alter our grouping to endorsement rather than outlet type, we find that remain outlets exhibit higher similarity between citizen group depictions across the board, but the differences are largest, again, for the EU citizen and ingroup comparison.
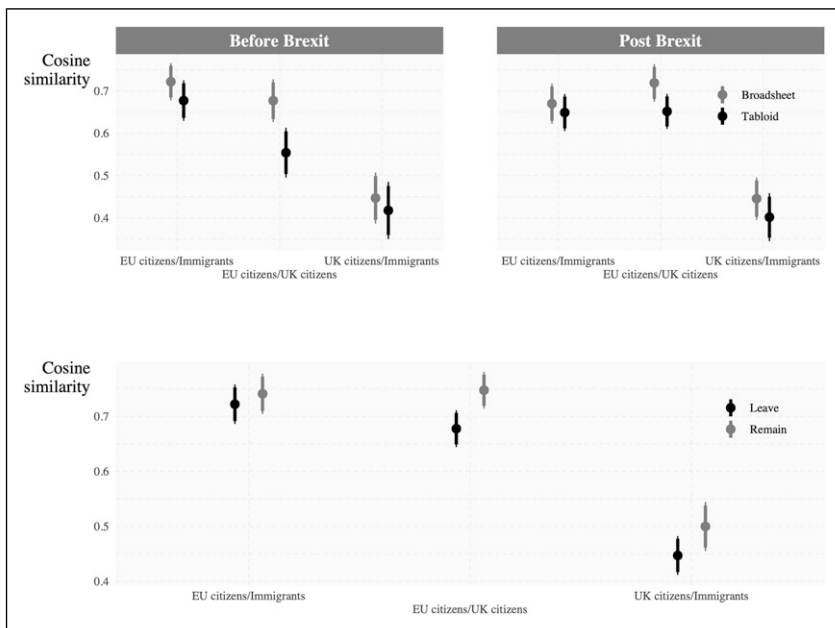
Figure 4. Campaign period and endorsement.

Discussion and conclusions

This study set out to examine whether and how references to different groups of citizens in the new coverage can create the image of in- and outgroups contributing to media-based othering. Specifically, we were interested in analyzing how the news media portrayed EU citizens in contrast to British citizens and non-EU immigrants in the context of the Brexit referendum. Our findings showed that EU citizens were portrayed in a very similar manner to immigrants – a well-established outgroup. This positions EU citizens as an outgroup that overwhelmingly shares features with immigrants, indicating homogeneity of representation. At the same time, EU citizens share more similarities with the ingroup of UK citizens than immigrants do. This puts them somewhere in the middle of the in/outgroup continuum which makes EU citizens a distinct group from UK citizens, but also from migrants. That both migrants and EU citizens were portrayed as dissimilar to UK citizens distances them from the ingroup. The portrayal of in- and outgroups in the news coverage thus confirms our expectations regarding the manifestation of media-based othering in the news media coverage.

With regards to differences between tabloids and broadsheets, we find that while both depict EU citizens as being very similar to immigrants, EU citizens share significantly more features with UK citizens in broadsheets than tabloids. This is in line with previous findings showing that broadsheets are more pro-European than tabloids (e.g., De Vreese et al., 2006). Our findings also suggest that media-based othering – though less

pronounced – not only applies to migrants but extends to citizens in other countries. Some have argued that the UK represents a special case compared to other EU member states, as its citizens are more afraid of Europeans than citizens from outside the EU (Freeden, 2017). It is however also possible that processes of globalization and Europeanization have led to foreign nationals, even those living in other countries, being perceived as a potential threat and thus being depicted as an outgroup in the news. This might be more pronounced for EU member states, where citizens have the opportunity to live and work in another member state. More research is needed that comparatively analyzes the in- and outgroup portrayal of different groups of citizens in different countries and contexts.

As always, these results come with limitations. First, there could be other formulations that make reference to various people, however capturing these in a comparative manner comes with quite some difficulties. Accordingly, our results should not be interpreted that these were all references to the people or various groups, rather a comparable subset of these potential mentions. Second, it could well be that other news articles made reference to immigration, for example, but they were not directly connected to Brexit headlines.

Third, while we investigated most similar terms resulting from the word embeddings and some expected associational relationships, the models are trained on data containing a limited set of features in comparison to other applications of similar methods. While this corpus is specific to our research interest, the feature set size comes as an important limitation in terms of the quality of the embeddings.

Overall, through this approach, we took an important step towards a framework that would accommodate the representation of foreign nationals beyond the distinctive category of migrants. Future research should explore further to what extent outgroup homogeneity features in the news coverage. Given preexisting, somewhat well-defined outgroups (such as immigrants), new groups may be linked to these by emphasizing shared features that place them on the in- and outgroup continuum. Rather than building specific portrayals from zero for a newly salient group, presentations can rely on what audiences might already be familiar with, through a transfer of properties. As anti-EU but also anti-immigrant sentiments are on the rise across Europe, a better understanding of how the media may contribute to othering and related phenomena such as populist sentiments is crucial.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ORCID iD

Stefanie Walter ⬛ https://orcid.org/0000-0002-9304-3469

Supplemental material

Supplement material for this article is available in online.

Notes

1. By that we mean mentions of migrants without any specific link to the EU or its member states, including the general use of the term migrant, but also mentions of migrants from geographic regions outside the EU or specific reference to non-EU migrants.
2. We keep online and offline versions of the same newspaper as separate outlets for the reported counts. This decision does not influence the results.
3. We prefer to start with a pre-defined list of terms and bi-grams to make sure they are relevant for our main question and are context specific for Brexit. Alternatively, named entity recognition could be employed, however, that would also require additional qualitative assessment.
4. We will refer to the global log-bilinear model as GloVe in the future, however, it is important to note that we do not use any pre-trained word embeddings. In all applications, we estimate the embeddings on our corpus.
5. In Supplementary Information 3, we present word usage differences for the word contexts. In Supplementary Information 4 we present our results in a different manner, by looking at which citizen group representation was most similar or different across the two outlet types.

References

Akkerman T (2011) Friend or foe? Right-wing populism and the popular press in Britain and the Netherlands. *Journalism* 12(8): 931–945.

Antoniak M and Mimno D (2018) Evaluating the stability of embedding-based word similarities. *Transactions of the Association for Computational Linguistics* 6: 107–119.

Atwell Seate A and Mastro D (2016) Media's influence on immigration attitudes: an intergroup threat theory approach. *Communication Monographs* 83(2): 194–213.

Bakarov A (2018) A survey of word embeddings evaluation methods, 09536. *arXiv:1801*.

Berry M, Garcia-Blanco I and Moore K (2015) Press Coverage of the Refugee and Migrant Crisis in the EU: A Content Analysis of Five European Countries Report prepared for the United Nations High Commission for Refugees (December 2015). http://www.unhcr.org/protection/operations/56bb369c9/press-coverage-refugee-

Blassnig S, Ernst N, Buchel F, et al.(2019) Populism in online election coverage. *Journalism Studies* 20(8): 1110–1129

Bleich E, Stonebraker H, Nisar H, et al. (2015) Media portrayals of minorities: Muslims in British newspaper headlines, 2001–2012. *Journal of Ethnic and Migration Studies* 41(6): 942–962.

Blinder S and Allen WL (2016a) Constructing immigrants: portrayals of migrant groups in British national newspapers, 2010–2012. *International Migration Review* 50(1): 3–40.

Blinder S and Allen W (2016b) *UK Public Opinion toward Immigration: Overall Attitudes and Level of Concern*. Oxford: The Migration Observatory.

Brüggemann M, Engesser S, Büchel F, et al. (2014) Hallin and Mancini revisited: Four empirical types of Western media systems. *Journal of Communication* 64(6): 1037–1065.

Conboy M (2006) *Tabloid Britain: Constructing a Community through Language*. London: Routledge.

Dathan M (2016) The True Cost of Our Open Borders Revealed: EU Migrants Are MORE Likely to Have a Job in the UK than British Citizens. *dailymail.co.uk*. Published on 7 June 2016. Available at: http://www.dailymail.co.uk/news/article-3628840/The-true-cost-open-borders-revealed-EU-migrants-likely-job-UK-British-citizens.html.

de Vreese CH, Banducci SA, Semetko HA, et al. (2006) The news coverage of the 2004 European parliamentary election campaign in 25 countries. *European Union Politics* 7(4): 477–504.

Dinas E and van Spanje J (2011) Crime story: the role of crime and immigration in the anti-immigration vote. *Electoral Studies* 30(4): 658–671.

Eberl JM, Heidenreich T, Herrero B, et al. (2018) The European media discourse on immigration and its effects: a literature review. *Annals of the International Communication Association* 42(3): 207–223.

Foner N (2015) Is Islam in western Europe like race in the United States? *Sociological Forum* 30(4): 885–899.

Freeden M (2017) After the Brexit referendum: revisiting populism as an ideology. *Journal of Political Ideologies* 22(1): 1–11.

Gabrielatos C and Baker P (2008) Fleeing, sneaking, flooding: a corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press, 1996-2005. *Journal of English Linguistics* 36(1): 5–38.

Garg N, Schiebinger L, Jurafsky D and Zou J (2018) Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences* 115(16): E3635–E3644.

Grove NJ and Zwi AB (2006) Our health and theirs: forced migration, othering, and public health. *Social Science & Medicine* 62(8): 1931–1942.

Hagendoorn L (1995) Intergroup biases in multiple group systems: the perception of ethnic hierarchies. *European Review of Social Psychology* 6(1): 199–228.

Hallin DC and Mancini P (2004) *Comparing Media Systems: Three Models of Media and Politics*. Cambridge: Cambridge University Press.

Harwood J and Roy A (2005) Social Identity Theory and Mass Communication Research. In: Harwood J and Giles H (eds). *Intergroup communication: Multiple perspectives*. New York: Peter Lang, 189–211.

Hogg MA (2006) Social Identity Theory. In: Burke PJ (ed). *Contemporary Social Psychological Theories*. Stanford: Stanford University Press, 111–136.

Innes AJ (2010) When the threatened become the threat: the construction of asylum seekers in British media narratives. *International Relations* 24(4): 456–477.

Kleinen-von Königslöw K (2012) Europe in crisis? Testing the stability and explanatory factors of the Europeanization of national public spheres. *International Communication Gazette* 74(5): 443–463.

Kozlowski AC, Taddy M and Evans JA (2019) The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review* 84(5): 905–949.

Krämer B (2014) Media populism: a conceptual clarification and some theses on its effects. *Communication Theory* 24(1): 42–60.

Kroon AC, Trilling D and van der Meer TG (2020) Clouded reality: News representations of culturally close and distant ethnic outgroups. *Communications* 45(s1): 744–764.

Levy DAL, Aslan B and Bironzo D (2016) *UK press coverage of the EU referendum*. Oxford: Reuters Institute for the Study of Journalism.

Mastro DE (2003) A social identity approach to understanding the impact of television messages. *Communication Monographs* 70(2): 98–113.

Matthews J and Brown AR (2012) Negatively shaping the asylum agenda? The representational strategy and impact of a tabloid news campaign. *Journalism* 13(6): 802–817.

Mazzoleni G (2008) Populism and the media. In: Albertazzi D and McDonnell D (eds), *Twenty-first Century Populism: The Spectre of Western European Democracy*. Basingstoke: Palgrave Macmillan, 49–64.

Mikolov T, Sutskever I, Chen K, et al. (2013) Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*: 3111–3119.

Moore M and Ramsay G (2017) *UK Media Coverage of the 2016 EU Referendum Campaign*. The Policy Institute at King's. Available at: https://doi.org/10.18742/pub01-029.

Nelson LK (2021) Leveraging the alignment between machine learning and intersectionality: using word embeddings to measure intersectional experiences of the nineteenth century US South. *Poetics* 88: 101539.

Ortiz M and Behm-Morawitz E (2015) Latinos' perceptions of intergroup relations in the United States: The cultivation of group-based attitudes and beliefs from English- and Spanish-language television. *Journal of Social Issues* 71(1): 90–105.

Pennington J, Socher R and Manning CD (2014) Glove: global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*: 1532–1543.

Poole E and Sandford E (2002) *Reporting Islam: Media Representations of British Muslims*. London: Tauris.

powell J (2017) *Us vs Them: The Sinister Techniques of 'Othering'–And How to Avoid Them*. London: The Guardian, Vol 8.

powell JA and Menendian S (2016) The problem of othering: towards inclusiveness and belonging. *Othering & Belonging* 1(1): 14–40.

Reinemann C, Aalberg T, Esser F, et al. (2016) Populist political communication: toward a model of its causes, forms, and effects. In: Aalberg T, et al. (eds) *Populist Political Communication in Europe*. New York, NY: Routledge, 12–28.

Rodriguez PL and Spirling A (2021) Word embeddings: what works, what doesn't, and how to tell the difference for applied research. *The Journal of Politics* 84(1): 101–115.

Rooduijn M (2014) The mesmerising message: the diffusion of populism in public debates in western. *Political Studies* 62(4): 726–744.

Share O (2018) Mobilizing migration: analyzing the role of the 'migrant' in the British Press during the EU referendum 2016 debate. In: Ridge-Newman A, León-Solís F and O'Donnell H (eds). *Reporting the Road to Brexit: International Media and the EU Referendum 2016*. London: Palgrave Macmillan, 25–46.

Startin N (2015) Have we reached a tipping point? The mainstreaming of Euroscepticism in the UK. *International Political Science Review* 36(3): 311–323.

Tong J and Zuo L (2019) Othering the European Union through constructing moral panics over 'im/migrant (s)'in the coverage of migration in three British newspapers, 2011–2016. *International Communication Gazette* 81(5): 445–469.

Walter S (2019) Better off without you? How the British media portrayed EU citizens in Brexit news. *The International Journal of Press/Politics* 24(2): 210–232.

Walter S (2017) *EU Citizens in the European Public Sphere*. New York, NY: Springer.

Author biographies

Stefanie Walter is an Emmy Noether research group leader at the Technical University of Munich's School of Social Science and Technology, Department of Governance. Her work focuses on the intersection between politics and communication, but also climate change and science communication. Her methodological focus in on quantitative methods and (automated) content analysis in particular. Her work has, e.g., in the International Journal of Press/Politics, European Political Science Review, and Information, Communication and Society.

Zoltán Fazekas is an Associate Professor of Business and Politics at the Department of International Economics, Government and Business, Copenhagen Business School (Denmark). His research is at the intersection of political psychology, political communication, and comparative politics and he mostly applies quantitative text analysis techniques and hierarchical models. His work has been published in the Journal of Communication, International Journal of Press/Politics, American Journal of Political Science, and Political Communication, among others.