Whose ideas are worth spreading? The representation of women and ethnic groups in TED talks

Accepted for publication at journal *Political Research Exchange* on July 12, 2019

Carsten Schwemmer*

University of Bamberg, Germany / Weizenbaum Institute for the Networked Society, Germany

Sebastian Jungkunz

University of Bamberg, Germany / Zeppelin University, Germany

Abstract

We investigate the representation of women and ethnic groups in TED talks, which reach a large online audience on YouTube with science-related content and topics on societal change. We argue that gaps in representation can create a misleading perception of science and the respective topics discussed in these talks. We validate annotations from an image recognition algorithm for identifying speaker ethnicity and gender to compile a data set of 2,333 TED talks and 1.2 million YouTube comments. Findings show that more than half of all talks were given by white male speakers. While the share of women increased over time, it is constantly low for non-white speakers. Topic modeling further shows that the share of talks addressing inequalities which affect both groups is low, but increasing over time. However, talks about inequalities and those given by female speakers receive substantially more negative sentiment on YouTube than others. Our findings highlight the importance of speaker and topic diversity on digital platforms to reduce stereotypes about scientists and science-related content.

Keywords: representation, women, ethnic groups, computational social science, YouTube, TED

*Corresponding author. E-mail: c.schwem2er@gmail.com

1 Introduction

With the digitization of human interactions affecting societies all over the world, digital platforms play an ever-increasing role for the creation and distribution of information. People consume information and entertainment content on social media platforms and via online news and videos, which heavily influences public opinion and shapes social identities. Similar to traditional media sources, it does not only matter what kind of information is distributed in the digital sphere, but also who is allowed to spread content. Despite the benefits of digitization, there is increasing concern about the level of digital democracy and whether audience concentration might lead to a new form of discourse elitism online. The important question therefore arises how disadvantaged groups across societies are affected by this development and to what extent they are represented in the digital sphere. While a large body of literature is available on representation and negative dispositions towards these groups in traditional media and entertainment outlets (Bleich, Bloemraad, and Graauw 2015; Shor et al. 2015), few research has been done with special focus on digital platforms.

In our paper, we investigate the representation of women and ethnic groups in TED talks. We focus on TED talks as a substantial part of the digital sphere because they reach a very large audience with science-related content and often discuss important matters of societal change. To do so, we build our theoretical argument on the literature about descriptive and substantive representation from the field of political science (Mansbridge 1999; Dovi 2002). A group is thus considered to be descriptively represented if a sufficient number of its members are part of a system of interest. In our case, this relates to what extent women and ethnic groups are allowed to give talks and represent themselves on a global stage like TED. In contrast, substantive representation is achieved when needs and issues of groups are addressed, e.g. when topics important for women and ethnic groups are discussed in TED talks. Gaps in both forms of representation can create a misleading perception of science and the respective topics that are discussed in these talks. Our first research question is therefore:

RQ1: To what extent are women and ethnic minorities descriptively and substantively represented in TED talks?

Furthermore, we examine how digital audiences respond to the representation of women and ethnic groups, as this allows us to gain insights about the public perception of these groups and the related topics. An important feature of most digital platforms are their responsive feedback systems, where users can give positive or negative feedback on content through digital interactions such as commenting, sharing, liking, or disliking. Unlike traditional newspaper articles, most online media are thus not consumed in isolation (D. Scheufele 2018). Comments from other users can have a decisive impact about what someone thinks about an article or video in the first place. Through these *social cues*, a neutrally framed article can, for instance, be perceived as heavily biased only due to viewers commenting behavior (Anderson et al. 2014). Such behavior has been observed across social groups and socio-demographic backgrounds, but highly social individuals in particular tend to be more affected by such developments and filter bubbles (Bar-Gill and Gandal 2017). On YouTube, positive or negative comments about TED talks and speakers could thus create polarization or amplify prejudices or stereotypes towards certain social groups or topics. Ultimately, such behavior can create feedback loops in which users adopt the predominating opinion of others (Rothschild and Malhotra 2014). Our second research questions is thus a more exploratory one, namely:

RQ2: How do descriptive and substantive representation of women and ethnic groups in TED talks affect viewer sentiment on YouTube?

To measure descriptive representation of speakers, we utilize and validate annotations of an image recognition algorithm. The algorithm detects faces within images of speakers and assigns probabilities for sex and ethnicity based upon physical appearance. Results of our analysis show that more than half of all TED talks were given by white male speakers and, while the share of talks by women increased from 2006 to 2017, it is constantly low for non-white speakers. To analyze substantive representation, we apply structural topic modeling on the transcripts of all TED talks to identify content in which needs and issues of women and ethnic minorities are addressed (Roberts et al. 2014). We identify a topic about inequality discussions which is relevant to both groups. Overall, this topic constitutes only three percent of talk transcripts, but the trend is increasing over time. Furthermore, sentiment analysis of YouTube comments suggests that the public sentiment is positive for non-white speakers, but negative for talks about inequalities, such as violence against women or racism. Talks given by women also receive substantially more negative and to some extent hateful comments than those given by men. We discuss how our findings highlight the importance of speaker diversity on global digital platforms to reduce stereotypes about science-related content.

2 Media effects and representation in the digital sphere

Representation in digital media such as TED talks is important as a large body of literature suggests that media consumption affects public opinion (see Tewksbury and D. A. Scheufele 2009; Valkenburg, Peter, and Walther 2016). In that regard, three theoretical concepts are particularly relevant (D. A. Scheufele and Tewksbury 2007): *Agenda-setting* describes how media influences the salience of certain topics in public discourse (McCombs and A. Reynolds 2009). *Priming* is considered as an extension of agenda-setting (D. A. Scheufele and Tewksbury 2007), which connects news content with certain benchmarks for evaluation that can result in a change of standards that people use to make assessments of particular topics (Iyengar and Kinder 1987). *Framing* on the other hand illustrates how public opinion is influenced through the way certain information is being presented. As such, frames define problems, they identify their causes, they render a moral judgment and suggest potential solutions for them (Entman 1993; D. A. Scheufele 1999).

Users consuming media in the digital sphere are thus influenced constantly through these mechanisms. For instance, if users of social media channels like Facebook are exposed to only a certain set of political information, it raises their perceived importance of these policy issues and it ultimately affects what people will think about (Feezell 2018). Priming of certain aspects by political elites in public debates, e.g. fake news, also influences citizens in their way of how they evaluate news media as a whole (van Duyn and Collier 2019). Priming can also subconsciously evoke certain feelings like aggressiveness (Buchanan 2015). Quite often it appears that biased content is most popular (Peer and Ksiazek 2011), as users become more engaged in user-content interaction when content is popular (Ksiazek, Peer, and Lessard 2016). There is also evidence for stronger user engagement in political campaigns when candidates tend to attack their opponents more often (Xenos, Macafee, and Pole 2017). And finally, framing not only shapes how we think about certain issues but also what people share with others on social media platforms (Valenzuela, Piña, and Ramírez 2017) or whether and how strongly they become engaged in political campaigns (Pond and Lewis 2019).

Media effects are subject to a multitude of issues like societal norms and values, pressures through interest groups or the ideological orientation of journalists and social media personalities (Tewksbury and D. A. Scheufele 2009). Related to that, a crucial factor is the question of who actually communicates information in the media. As Merskin (2017, p. 1098) argues, mainstream media "are powerful hegemonic tools" that enforce cultural reproduction of dominant ideologies e.g. in terms of race and gender, but also regarding economic or social values. If then representation of certain societal groups in the media is either unbalanced, inaccurate or stereotypical, it not only negatively shapes the perception of how non-minority members think about such groups and their interests, but also how these groups themselves think about their own group. It is therefore both the visibility and the nature of the visibility of groups in the media that matters about how we perceive certain phenomena in society (Merskin 2017).

When it comes to the representation of disadvantaged groups in traditional media outlets, e.g. in newspapers and television, a large number of studies already investigated outcomes for women and ethnic minorities. For instance, scholars analyzed media coverage of migrants and minorities to see what kind of information is presented, who is allowed to participate and how minorities are represented (Bleich, Bloemraad, and Graauw 2015; J. Miller 2006). Many similar studies have been conducted about the representation of women, e.g. in radio, television and film (Fonda, Morgan, and Steinem 2017) as well as in newspapers (Harp, Bachmann, and Loke 2014) and other printed news (Shor et al. 2015). The overall image depicted by studies about these groups is clear: both women and certain ethnic groups often suffer from under-representation and coverage that enforces stereotypes and negative attitudes.

Scholars also started to examine whether similar patterns exist for the increasingly available share of digital content, but the body of available work is still small in comparison. Jia et al. (2016) collected visual and textual data of online newspapers to analyze how gender is represented in online news. They found that women were more likely to be represented visually than they were mentioned within texts and that online news sources are still male-dominated. Other scholars found that social media platforms, such as YouTube (Guo and Harlow 2014), Facebook and Twitter (Matamoros-Fernández 2017) are used to spread content enforcing stereotypes or xenophobic attitudes about ethnic minorities. However, while these studies provided important insights, they do not enhance our knowledge about the connection between representation and the interaction with the digital sphere, e.g. in comment sections of blogs, web pages, or social media platforms. Qualitative work by Sobieraj (2017) about digital sexism shows that studying these interactive environments is crucial to make sense of how the public perceives digital content. This interactive part of the digital space enables users to give positive or negative feedback on content by digital interactions such as commenting, sharing, liking, or disliking. One key contribution of this paper is thus not only to study representation in the digital sphere (RQ1), but also the public sentiment to it (RQ2).

For our theoretical concept of *representation*, we draw upon terminologies from the large body of political science literature about representation in political systems. A group is thus considered to be *descriptively represented* if a sufficient number of its members are part of a system of interest (Pitkin 1967; Mansbridge 1999; Dovi 2002). As a simple example, the gender distribution in most societies is close to even, which makes it desirable that half of all members of national parliaments are women. In comparison to representative political systems, it is often difficult to evaluate whether the descriptive representation of certain groups is adequate in the digital sphere. Political systems usually have well-defined populations of people for which they are responsible, such as all residents of a country. For digital platforms, this is not generally true. Web pages like that of TED media or social media platforms like YouTube reach global audiences and distributions of their socio-demographic attributes are often unknown. Thus, even though it is important to what extent women and ethnic groups are allowed to give talks and represent themselves on a global stage like TED, it is also difficult to agree upon what constitutes *enough* representation. Nevertheless, members of disadvantaged groups need to be represented in the digital sphere, as under-representation can lead to the same issues that scholars found in traditional media sources.

Moreover, *substantive representation* assesses system responsiveness, a normative ideal of democracy. For instance, in an electorate with a high share of citizens of immigrant origin, their needs and interests should find more consideration in the activities of their representatives (Dahl 1971). In our case, this applies when topics that are important for women and ethnic groups are discussed in TED talks. TED speakers are no elected representatives, but it is still crucial whether they address the needs and interests of women and ethnic groups on the TED stage. In that regard, the substantive representation of groups can be connected to and affected by descriptive representation, but not necessarily so. Often, it is believed that a "critical mass" or "critical actors" can be sufficient (MacDonald and O'Brien 2011; see also Celis and Childs 2008; Celis and Erzeel 2015). Group members are more likely to be substantively represented by people that *look alike* and have similar needs. In this paper, we make use of both concepts for representation and apply them to studies of the digital sphere. We refer to descriptive representation as the presence of women and ethnic groups on the TED stage as speakers. As for substantive representation, we examine whether needs and issues relevant to women and ethnic minorities are addressed in the content of TED talks. Being invited for a TED talk gives speakers the possibility to spread content that is important for them and we therefore assume that descriptive representation in the digital sphere also fosters substantive representation.

Much in the way of the *thermostat model* (Wlezien 1995, Soroka and Wlezien 2010), public sentiment is likely to react according to how well certain groups of society are represented both in terms of their physical presence and the topics that concern them the most. As this relationship is heavily amplified through media attention (Williams and Schoonvelde 2018), we will use an exploratory approach to investigate reactions towards descriptive and substantive representation. It is possible that stronger representation of certain ethnic groups can also increase negative sentiments among some members of society. For instance, women and ethnic minorities often suffer from hate speech or other forms of discrimination in online environments and digital media outlets (Chetty and Alathur 2018). Thus, it is likely that we can observe similar patterns for TED Talks, too. In general, we expect that stronger representation of women and ethnic groups will also trigger reactions in public sentiment.

3 Digital democracy and discourse elitism

Based on the premise that "meaningful democratic participation requires that the voices of citizens in politics to be clear, loud, and equal" (Verba, Schlozman, and Brady 1995), many expected the Internet to become a powerful tool to democratize communication and politics, an "army of Davids" against big media and political elites (G. H. Reynolds 2006). The Internet, so the argument of many early journalists, commentators and scholars, would empower citizens to participate in online communities that create content collaboratively, thereby diminishing the influence of traditional media juggernauts (Benkler 2006). In fact, there is evidence that the Internet has increased the potential for citizens to make meaningful and widely heard contributions online, or that so called "burglar alarm" models of reporting misdemeanor and wrongdoing become more effective (Hindman 2008; Zaller 2003). However, there is also growing concern about the online establishment of new forms of discourse elitism.

According to Matthew Hindman, the infrastructure of the web does not decrease centralization of news content, but often it actually increases centralization to levels way beyond traditional news outlets in the offline world (Hindman 2008; Hindman 2009; Hindman 2018). This is true for four reasons: first, online traffic and the link structure of the internet follow a power-law distribution, in which few sites receive the bulk of visits and most sites remain fairly untouched (Huberman et al. 1998). Thus, most people get to see similar content and the potential of immensely diversified information is never reached. Much of this has to do with the way citizens search for information online. Secondly and against popular conviction, even digital natives are not universally savvy online (Hargittai 2010) and many people rely on well-known and longestablished procedures to browse the web (Hargittai et al. 2010). Such skill-based habits of use can then induce cognitive lock-in, e.g. that the repeated use of products or sites online creates cognitive switching costs which increasingly prohibit the user from looking for potential alternatives (Murray and Häubl 2007; Johnson, Bellman, and Lohse 2003). Thirdly, "digital distribution is never free" and "costs of audience building" are huge (Hindman 2018, p. 167). Unlike traditional media, it is not the production, but rather the distribution of information that is so expensive. Digital survival hinges upon one's stickiness, e.g. the ability to attract users and make them return to your website over and over again (Hindman 2018). Large online players have therefore a considerable advantage, as they can rely on economies of scale in terms of staff, equipment, speed and especially data that can be used for personalization through recommender systems (Sundar et al. 2015; Valkenburg, Peter, and Walther 2016). That way, large sites can direct users much faster to their desired content than any small niche site could do it. And finally, those who get heard in the online sphere are not representative for average citizens, but to a substantial degree "well-educated white male professionals" (Hindman 2009, p. 128). For the US, it has been shown that most successful blogs are run by educational, business and technical elites or journalists from traditional media outlets with Ivy League degrees (Hindman 2009). Overall, digital activism is strongly affected by ethnicity and class (Schradie 2018). Entman and Usher (2018) further extend this line of thought by stressing the importance of platforms, algorithms, digital analytics, ideological media and rogue actors in framing news content and ultimately shaping public opinion.

Such developments can become alarming if the content that is distributed through a small number of big players represents the opinion of a particular interest group only. Spreading lopsided arguments to large audiences can tilt public opinion in one way or the other. For instance, if a news outlet predominantly hosts neo-liberal economists and politicians to discuss major topics like the Euro crisis or the refugee crisis, some solutions might be presented as if there is no viable alternative. User and viewer opinion could thus be pushed into a certain direction, whereas commenting behavior can add to that effect and increase polarization among viewers. Unfortunately, the growing presence and consumption of *soft news*, which present news in episodic frames without additional context and often in personalized manners, pulls viewers' opinions heavily towards one side or the other (Baum 2004; Boukes and Boomgaarden 2015; Ter Wal 2002). Thus, it is particularly important to study representation of disadvantaged groups like women and ethnic minorities in successful online media outlets such as TED talks. Furthermore, it is relevant to examine how topics relating to these social groups are represented in the online sphere and how viewer respond to such topics.

4 Why study TED Talks?

Studying representation of social groups in digital media is empirically challenging due to a huge number of potentially relevant data sources. Depending on research motivations, this can make it difficult to assess reasons for choosing a particular source over another. We consider TED talks as a particularly interesting case, because they introduce science-related content to the public and in doing so they reach a very large online audience. TED talks are part of conferences organized by the media organization *TED: Technology, Entertainment and Design*, which was founded in 1984. Talks of TED conferences are distributed for free across several platforms under the slogan *Ideas Worth Spreading.* Invited speakers were scientists like Stephen Hawking and entrepreneurs like Bill Gates, as well as some activists and entertainers. In general, the populist nature of TED talks (Tsou et al. 2014) can be characterized by a sales pitch atmosphere, passionate styles of delivery and hints at feelings of self-actualization and inspiration (Ludewig 2017). For this reason, some scholars consider TED talks as a source of information for the masses rather than for scientists (Sugimoto and Thelwall 2013). Speakers giving TED talks also present relatively few counter-arguments to their proposed ideas and thus potentially undermine alternative viewpoints (Singh Chawla 2016). Moreover, TED conferences are often considered as elitist events, restricted to those who can pay significant entrance fees. To give one example, attending a 2018 TED conference in person was priced at around $10,000, which results in a very special audience (Schwartz 2018; Turnaround Management Association 2017). While potentially any person can be suggested as a speaker via a nomination form on the TED website, it is unknown how TED exactly selects the speakers for its elitist stage.

One reason why we analyze TED talks in this paper is the significant size of their audience. Talks are available at the TED homepage, via several mobile applications and most importantly the YouTube video sharing platform. In 2012, TED celebrated one billion views for videos uploaded on its own web page, stating that talks were being viewed at 1.5 million times a day. In April 2018, Socialblade Analytics ranked TED's YouTube channel at top 250 global, as it attracted more than nine million subscribers and received more than one billion views in total.

TED talks also get covered regularly in global publications like *The Guardian* (Giussani 2015) and *The New York Times Magazine* (Dominus 2017). The predominant language of talk content is English, but transcripts are available for more than 100 languages. For most YouTube videos of talks, captions are also available in several languages. Regarding scientific coverage, only few studies investigated TED talks. One study examined presentation characteristics of TED talks and concluded that they can be a valuable source for teaching and communicating ideas to students (Kedrowicz and J. L. Taylor 2016). Other scholars found that women gave fewer TED talks than men (Sugimoto, Thelwall, et al. 2013). In sum, TED talks reach a very large audience and provide content in many languages, which makes them a relevant case for analyzing representation in the digital sphere.

Another reason for focusing on TED talks is that they are considered as a "highly successful disseminator of science-related videos" (Sugimoto, Thelwall, et al. 2013, p. 1) and the content of talks is often related to important matters for societal change. Their popularization of science and digital content about *ideas worth spreading* draws attention from people across several societies. Due to the large audience, under-representation of certain groups on the TED stage is particularly problematic. If, for instance, TED talks would predominantly be presented by white male speakers, this could lead to the development or amplification of stereotypes among the audience. Decades of *Draw-A-Scientist* studies have shown that as children get older and stereotypes manifest, they predominantly associate science with men (Chambers 1983; D. I. Miller et al. 2018). Other work by scholars further suggests that the development of stereotypes (Shor et al. 2015), negative attitudes (Van Klingeren et al. 2015) and lowered self-esteem (Martins and Harrison 2012) can at least in part be ascribed to media exposure. It is therefore important for the global audience of TED that women and ethnic groups are adequately represented descriptively and substantively to not further enhance stereotypes and negative attitudes.

5 Data and methods

To analyze the representation in TED talks, we first applied web scraping techniques to retrieve content about talks and corresponding speakers from the TED homepage. Content from the TED website is distributed as creative commons licensed material. We collected information for the very first talk given in 2006 up until the time of data collection in May 2017. Second, we used YouTube Data Tools (Rieder 2015) to find the related YouTube videos for these talks and collected meta data about video metrics such as dislikes and video views. Subsequently, we also retrieved all 1.2 million available user comments for the YouTube videos and used a combination of approximate string matching and manual coding to merge data from TED and YouTube. In this process, we noticed that if talks appear on the TED page, they will generally also be available on the main YouTube channel, regardless of whether they are conventional talks, from the TEDx subbranch, or from specific events like TEDWomen. Our sample thus includes all talks that TED considers to be relevant enough for its main channel. Furthermore, we restricted our sample by removing a small number of talks that do not contain language (e.g., music concerts or dancing) and talks without human speakers as main presenters (e.g. an entertainment talk about Einstein the Parrot).

5.1 Descriptive representation: speaker gender and ethnicity

Capturing attributes of thousands of speakers by manual coding is a resource draining task. We therefore decided to utilize an automated approach in form of an image recognition algorithm. We discuss ethical considerations of using image recognition systems for research purposes in our concluding section. As facial recognition software is increasingly deployed by companies and government agencies (Dearden 2018), scholars have recently started to analyze their performance. They found that face recognition algorithms by Microsoft, IBM and Face++ can produce substantially biased results by identifying the sex of darker-skinned females less accurately in comparison to other skin tones (Buolamwini and Gebru 2018). In this work, we use the Kairos diversity recognition software for gender and ethnicity annotations. In 2018, the company released a statement about restricting access to these features in light of concerns about face recognition systems and their consequences for privacy and safety (Brackeen 2018). As of February 2019, gender and ethnicity annotations are still available via the Kairos Application Programming Interface (API). The Kairos API detects faces within images of speakers and assigns corresponding probabilities for sex and ethnicity based upon facial features. Our data set contains images of all speakers publicly available at the TED homepage, which we used as input for the diversity recognition algorithm. Figure 1 shows the annotation output for two publicly available example images of TED speakers.

Depicted on the left-hand side is Kimblerlé Williams Crenshaw, a scientist and full-time professor who gave a TED talk about race and gender bias. The famous entrepreneur Bill Gates is a regular TED speaker and depicted on the right-hand side. In both cases, the algorithm correctly annotates the sex of the speakers. Regarding ethnicity, it is only possible to differentiate between Asian, Black, Hispanic, White and a category for ethnic groups other than those mentioned. In both cases, the mode prediction of the algorithm is in line with labels assigned by human coders (see S1). We automated the process of sending each image to the image recognition service and

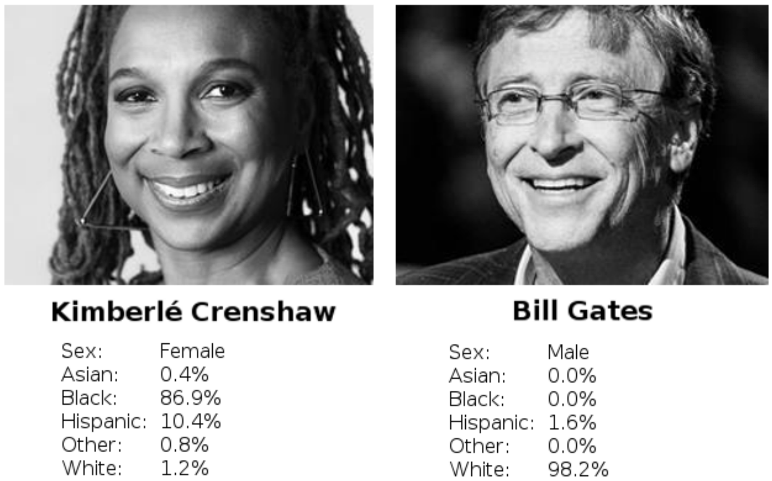

Fig 1. Example outputs of diversity recognition algorithm.

then merged all annotations with our data. For ethnicity, we always coded the mode prediction of the algorithm as final value. As can be seen in supporting information S1, the majority of the probability mass for each ethnic group lies at either 0 or 1, which is why using the mode does not result in a noticeable information loss.

To evaluate algorithmic performance, we conducted several validity checks and computed inter-rater agreements between two human coders and the algorithm for 200 randomly selected images of TED speakers. For all images in this sample, we computed Fleiss' Kappa for pairwise and overall ratings between two human coders and the algorithm (see S1). In total, Kappa values are satisfying, with 0.85 for ethnicity and 0.95 for sex. For the ethnicity categories *Hispanic* and *Other*, agreement is lower in comparison to the remaining categories. Qualitative inspections revealed that the small number of disagreements for these categories can not solely be ascribed to poor algorithmic performance as both, pairwise comparisons between human coders as well as between humans and the algorithm were not always in line. For instance, some speaker images were labeled as White by one rater and Hispanic by another rater or the algorithm. Overall, our validation results nevertheless suggest that annotations from the image recognition algorithm are sufficient for our research task. We also conducted additional analyses for which we treated human annotations as the gold standard. In doing so, we assessed the performance for predicting ethnicity and gender with F1 scores and compared predictions of the image recognition algorithm with predictions based on the names of speakers (Wais 2016; Imai and Khanna 2016). Results indicate that the image recognition algorithm outperforms name-based approaches for both ethnicity and gender (see S1).

One issue with our data is that a small number of TED talks were given by groups and not by single persons, which makes it difficult to analyze descriptive representation when attributes like ethnicity are not identical across group members. We therefore identified talks with more than one speaker and only included these talks in our data if all members have identical annotations for sex and ethnicity, which reduces our final data set to 2,333 TED talks.

Regarding descriptive representation, we need to highlight one important limitation at this point: algorithms such as the image recognition algorithm obviously neglect the fact that gender is a non-binary construct. What kind of labels an algorithm assigns may not be in line with the gender a person identifies herself with. The same applies for race or ethnic groups, which are also not objective categories, but rather social constructs that may vary across cultures. With that in mind, we still believe that our (over-) simplified indicators are useful to examine the representation of women and ethnic groups.

5.2 Substantive representation: content of TED talks

Ideally, as TED talks are available as videos, substantive representation of women and ethnic groups could be analyzed in both, visual and audio data. Scholars have recently started to work on signal processing for audio data (Knox and Lucas 2018), but models for video content that are useful for social scientists are still to be developed. For this reason, we focus on the transcripts of TED talks to examine substantive representation, as we assume that what speakers say in their presentations is the most important way to address the needs and issues of disadvantaged groups. We chose the English language because TED talks are almost exclusively given in English and it is also the language for which the most transcriptions are available.

To analyze substantive representation we apply topic modeling on all talk transcripts. Topic models are a method for automated content analysis. They allow to discover latent themes from text documents, where a topic can be understood as a set of words representing these themes and documents are represented as mixtures of topics. To provide an example for this paper, after fitting a topic model, a talk transcript might contain content related to a topic about *technology*, consisting of words like *computer, machine, device* and *algorithm* with a proportion of 60%. In addition, 30% of its content could capture a topic *internet* with words like *internet, online, website, link* and the remaining 10% would include other topics. To prepare our textual data for topic modeling, we used the programming language *R* (R Core Team 2018) and the corresponding packages *quanteda* (Benoit 2018) and *tidyverse* (Wickham 2016) to process all transcriptions into a corpus with common methods of text analysis. Texts were treated as bags of words in which each term represents a feature and word order is ignored. In addition, terms without semantic meaning, such as words like *and* or *the* and very infrequent terms were removed from the corpus. We further applied a stemming algorithm to all terms, so that words with similar semantic meaning, such as *inequality* and *inequalities*, get reduced to their common word stem *inequ* (Grimmer and Stewart 2012).

As we expect that substantive representation of disadvantaged groups can be affected by descriptive representation and developments over time, we utilize a novel variant of topic models called the Structural Topic Model (Roberts et al. 2014). Structural Topic Models enable us to not only examine topic proportions for each TED talk, but also to analyze how these proportions vary dependent on the date of the TED talk, speaker gender and speaker ethnicity. Although topic modeling is very useful to automatically categorize large text corpora, one limitation is that the number of topics has to be determined by the analyst. To find the best model for our research goals, we utilized the R package *stminsights* (Schwemmer 2018) to qualitatively inspect and validate several models. More details about our validation procedure are available in the Supporting Information S2. The validation showed a model with 30 topics to be superior to others, both in terms of desirable statistical properties as well as its usefulness for our research task. We therefore chose this version as our final model and assigned labels to each of its topics.

5.3 Public sentiment on YouTube

To examine our second research question - whether descriptive and substantive representation affect viewer sentiment - we (a) apply sentiment analysis on all 1.2 million YouTube comments of TED talks and (b) count the respective (dis-) likes of each video. Sentiment analysis describes methods to measure people's opinions, sentiments, attitudes and emotions from language (Liu 2012). As it pertains to all sentiment analyses (e.g. Tsou et al. 2014), our results apply to those who comment on YouTube, which means we cannot infer on the sentiment of people who remain silent. Nevertheless, YouTube comments appear below the corresponding videos and therefore visible to viewer regardless of whether they are commenting themselves. It is therefore reasonable to assume that the comment sentiment of TED talks is likely to affect many viewers that are not sharing their opinions on the platform. While there are different ways to conduct sentiment analysis, most of them rely on the use of dictionaries. With the aim to detect whether people speak in a positive or negative way about something, polarity dictionaries usually include words and corresponding weights. These dictionaries are then used to produce an overall sentiment value for each text document. Although sentiment analysis is a commonly used instrument for analyzing social media discourse, it is associated with a number of limitations (Puschmann and Powell 2018). Among other issues, sentiment analysis is constrained by two problems in particular: first, results strongly depend on the dictionary in use. A dictionary created by researchers for examining moods in short social media messages like Tweets is unlikely to properly capture the mood in, for instance, political speeches. Second, most implementations for sentiment analysis are very simple and do not account for basic features of human language like valence shifters. Valence shifters like *not* ("I do not like it") or *really* ("I really like it") are commonly used in spoken and written language to alter the original meaning or sentiment of words.

In our paper, we apply a novel, sentence-based variant of sentiment analysis from the *sentimentr* package (Rinker 2018) that combines established dictionaries for polarity terms (Jockers 2017; Hu and Liu 2004) and valence shifters. Moreover, this implementation also accounts for the use of internet slang and emoticons, which frequently occur in social media data. For each TED talk, we calculate the average sentiment value over all corresponding YouTube comments. 13 YouTube videos of TED talks were not included in the analysis, as comments on these videos either did not yet exist or were disabled on the platform. Overall, the average comment sentiment across all TED talks is at 0.06 and the median is at 0.02. The following examples show two YouTube comments with very positive or negative sentiment values:

- "epic epic epic epic epic epic epic epic epic epic epic epic" (sentiment score: 2.77)
- "wow, sir I am really really inspired by your speech every thing what you explained was awesome but the last minutes of your speech in this video was pure motivation 'Yes We Can' that's brilliant." (sentiment score: 2.96)
- "GO FUCK YOURSELF YOU ARROGANT PRICK GO FUCK YOURSELF YOU ARROGANT PRICK GO FUCK YOURSELF YOU ARROGANT PRICK" (sentiment score: -3.74)
- "lies, lies & more lies" (sentiment score: -2.30)

After calculating sentiment scores for each video, we use these scores as dependent variable in a generalized linear model. As independent variables, we incorporate speaker gender and ethnicity and the topic proportions from our structural topic model, while controlling for popularity (views on YouTube) and time (YouTube upload date). At last, we calculated predicted values for our representation covariates with the R package *ggeffects* (Lüdecke 2018). Given the limitations of sentiment analyses, we computed another regression model with the same covariates for predicting the number of dislikes for each video. Model information is available in the Supplementary Information S4 and the results are very similar to those obtained from sentiment analysis.

6 Results

6.1 Descriptive representation

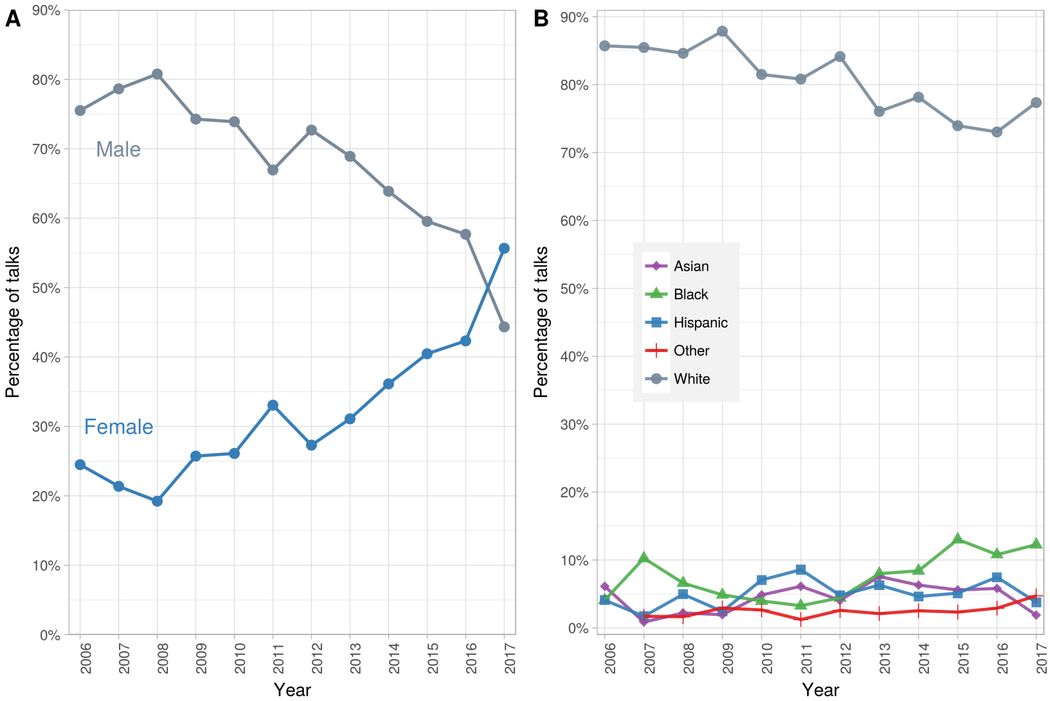

Fig 2. Percentage of TED talks over time. Fig 2A: percentage by speaker gender. Figure 2B: percentage by speaker ethnicity.

Applying the image recognition algorithm to the images of all TED speakers, we find that overall, 68% of speakers are men and thus only about one third of all speakers are women. With regards to ethnicity, 80.2% of all speakers are classified as white, with the remainder consisting of 7.3% Black, 5.4% Hispanic, 4.7% Asian and 2.4% other ethnic groups. These numbers show that male speakers are substantially more often allowed to present at TED stages than women. Moreover, only one out of five speakers' ethnicity is non-white. To put this into another perspective, combining both attributes results in 56.2% percent of speakers being both white and male.

Figure 2 further illustrates how descriptive representation for women and ethnic groups developed over time. Regarding speaker gender, the figure shows that, despite the overall low share of women speakers on TED stages, the trend increased over the years. In fact, in the year of our data collection up until May 2017, women were for the first time more often present as speakers than men. However, when it comes to ethnicity, the representation of non-white speakers is constantly low across time, with the share of white speakers declining only by a small margin. To put this in context, we can compare numbers for TED talks given between January 2016 and May 2017 with demographics from the United States, where a TED headquarter is located and many TED conferences took place. As can be seen in Table 1, TED improved in terms of an equal distribution of sexes.

Table 1. Percentages of gender and ethnic groups in the United States (2017) and TED Talks (Jan. 2016 - May 2017)

	US	**TED Talks**
White	61.5	74.5
Black	12.3	11.3
Asian	5.3	4.3
Hispanic	17.6	6.4
Other	3.4	3.5
Male	49.2	52.8
Female	50.8	47.2

Source: U.S. Census Bureau 2019, own calculations for TED talks

At the same time, white speakers are over-represented and some ethnic groups like Asians and especially Hispanics are still substantially under-represented in comparison to the United States. A much better comparison would be possible with knowledge about the geographical distribution of TED viewers, for instance on YouTube. Unfortunately, TED did not provide such data after a request from the authors. Nevertheless, it can be expected that the share of non-white viewers is higher in many countries outside North America and Europe. For this reason, speakers with non-white ethnic groups are still under-represented in TED talks.

6.2 Substantive representation

We identified one out of 30 topics from our structural topic model to be strongly related to the substantive representation of women and ethnic groups. We labeled this topic as *inequality* topic, as the corresponding talks predominantly address unequal treatment of either women or ethnic groups. The topic accounts for about 3% of all TED talk transcripts. Table 2 shows the most important stemmed terms as indicated by the *frex* metric, which captures terms that are both frequent and exclusive for a given topic (Lucas et al. 2015, p. 19). Labels, proportions and frex terms for all other topics are included in Supporting Information S2.

Table 2. Frex terms and titles of representative talks for *inequality* topic.

Terms	Talk titles
women, men, gender, gay, sexual, sex, girl, woman, rape, feminist, femal, black, male, boy, slaveri, marri, equal, abus, violenc	1. Violence against women - it's a men's issue 2. What does my headscarf mean to you? 3. How to overcome our biases? Walk boldly toward them 4. The racial politics of time 5. The fight against sex slavery 6. A warrior's cry against child marriage

As indicated by the frex terms, the topic captures content about gender (*women, men, gender, girl, boy, feminist*) and sexuality (*gay, sexual, sex*). References to ethnicity (*black*) and several terms related to violence and misuse of power (*rape, slaveri, abus, violenc*) are also apparent. The stemmed term *equal* refers to *equality*, suggesting another important aspect of substantive representation for women and ethnic groups. In addition, we utilized a feature of the structural topic model to find the most representative TED talks, for which corresponding titles are also included in the table. The titles provide further evidence that the topic is strongly related to inequality structures, bias and negative attitudes towards women and certain ethnic groups, but also children.

As topics of a structural topic model are by design allowed to correlate with each other, we can examine which topics frequently co-occur together in TED talks. To visualize topic connections, Figure 3 contains output from hierarchical ward clustering of topic proportions on the left-hand side.

The topic clusters reveal that inequality is correlated with content about family as well as children & school, but also to a lesser extent with the topics law & politics, war & terror,

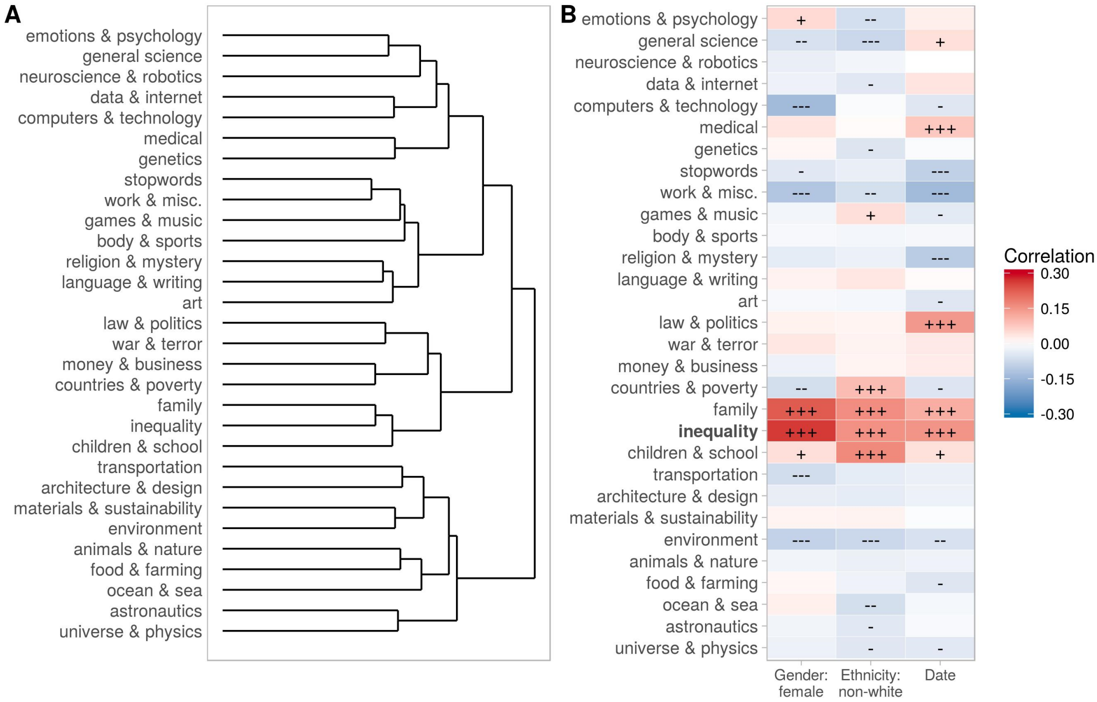

Fig 3. Topic clusters and correlations. F3A: hierarchical ward clustering of topics. Fig 3B: pearson correlations between prevalence variables of the structural topic model and proportions of all topics. Characters denote p values for positive (+) and negative (-) correlations, adjusted for multiple comparisons ($3 = p < 0.001$, $2 = p < 0.01$, $1 = p < 0.05$).

money & business and countries & poverty. As we incorporated speaker gender and ethnicity in the estimation process of our topic model, we can assess whether these covariates affect the likelihood to talk about inequalities related to women and ethnic groups. We also incorporated the date of talks to examine whether TED preferences for certain topics change over time. Figure 3B includes correlations between all topics and covariates with p values adjusted for multiple comparisons. As expected, descriptive representation affects substantive representation and women as well as non-white speakers are more likely to discuss corresponding inequalities on the TED stage. Supporting Information S3 contains estimates for topic proportions dependent on the same covariates, which are identical to the correlation patterns. The output from Subfigure 3B further shows that over time, TED talks increasingly covered content related to inequality. It can also be seen that women are less likely to give talks about computers & technology and that the environment topic is predominantly discussed by white male speakers. In summary, inequalities relevant for women and certain ethnic groups are addressed in a small but increasing share of talks, suggesting that both groups are substantively represented on the TED stage. Moreover, they are also more likely to talk about these inequalities.

6.3 YouTube sentiment

To recap our final research question, we examine whether descriptive and substantive representation of women and ethnic groups affects how TED talks are perceived by the digital audience. Figure 4 illustrates regression estimates for the average sentiment of YouTube comments for each TED talk. Topic proportions of topic models always sum up to 1, which is why including every topic proportion in regression models induces perfect collinearity. For the following analysis, we therefore removed one topic that is irrelevant for our research task, labeled as work & misc, from our regression model.

Figure 4A shows standardized coefficients for speaker attributes, topic proportions and controls. Following the recommendation by Gelman (2008) numeric variables were divided by two standard deviations, so that they can be compared to binary variables for gender and ethnicity. The output suggests that, surprisingly, while holding speaker gender and the topical content of TED talks constant, the public sentiment of non-white speakers is positive and they receive more positive comments than white speakers. Supplementary material S3 includes a visualization from an additional regression model including an interaction between ethnicity and date, showing that sentiment differences between white and non-white diminish over time.

With regards to speaker gender, the sentiment of talks given by female speakers is generally more negative in comparison to men on the TED stage. However, the most important predictors for receiving negative feedback on YouTube are related to talk content. Average sentiment values of talks are most negative for high topic proportions of the substantive representation topic (inequality) and for talks about law & politics and war & terror. This could be further evidence for the hostile media phenomenon, e.g. the perception of people with strong pre-existing attitudes that media coverage is biased against their own point of view (Gunther and Chia 2001). As a result, sentiment values in such highly polarizing talks are also substantially more negative. In addition to standardized coefficients, Figures 4 B.1-B.3 show effect estimates for representation covariates, where all other variables where held at their observed values. The output confirms that descriptive representation is related to public sentiment, but the effect of substantive representation is stronger. To test the robustness of our results, we also analyzed

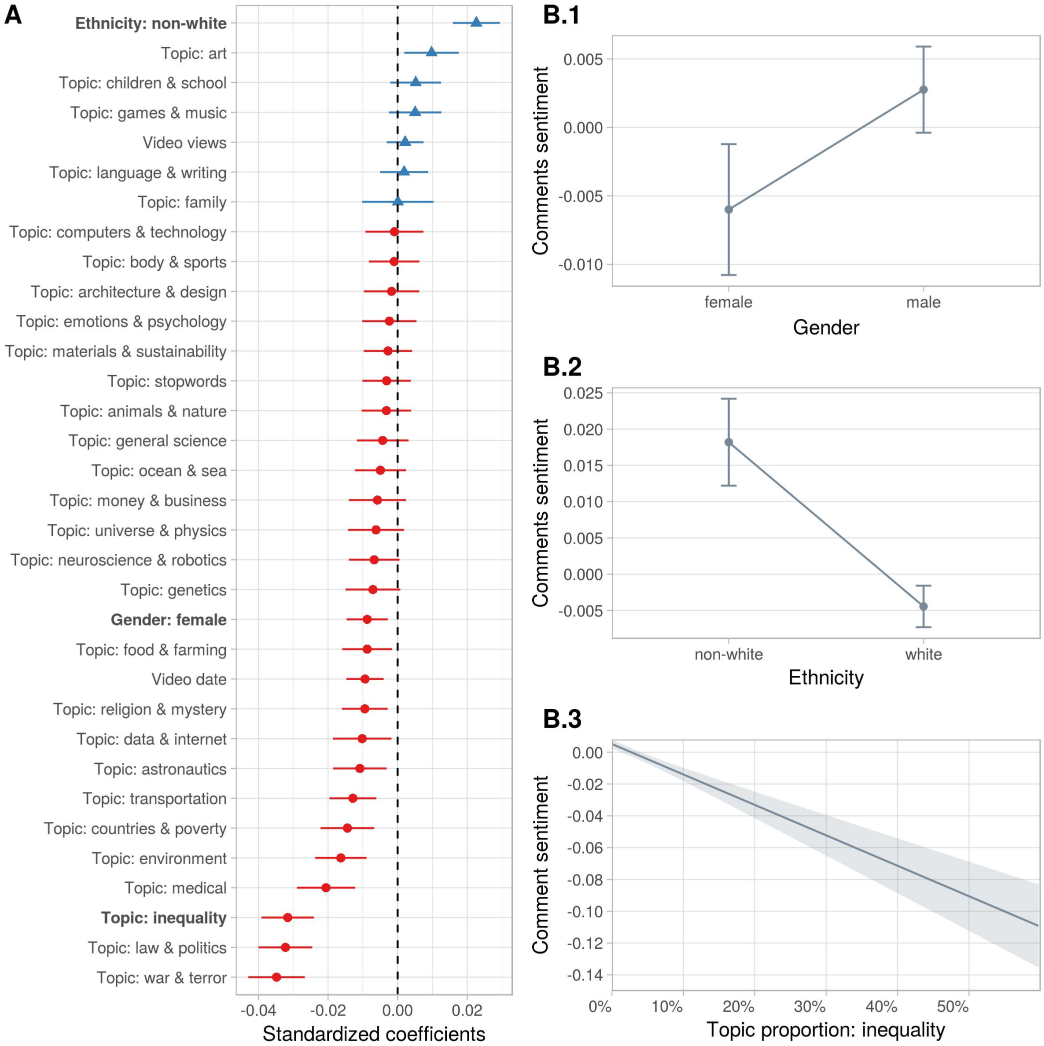

Fig 4. Regression estimates for the average comment sentiment of TED YouTube videos. Figure 4A: standardized coefficients of generalized linear model with 95% confidence intervals. Figures 4B.1 - 4B.3: predicted sentiment values for descriptive and substantive representation covariates.

the number of dislikes a TED talk on YouTube receives, using a negative binomial model and the same set of covariates. Regression tables and visualizations for both models are available in Supporting Information S3. Results of the dislike regression model are very similar, also showing a positive sentiment of talks by non-white speakers, but negative sentiments for female speakers and talks about inequality.

7 Discussion and conclusion

This paper was motivated by our interest in the representation of women and different ethnic groups in the digital sphere. By utilizing automated methods for image annotation and text analysis, we examined to what extent members of both groups give TED talks and whether their specific needs are discussed on stage. We showed that more than half of all talks were given by white male speakers and, while the share of talks by women increased over time, it is constantly low for non-white speakers. We further identified a small but increasing share of TED talk content being strongly related to inequalities and the substantive representation of women and ethnic groups. Both women and certain ethnic groups were more likely to discuss such inequalities on stage. Moreover, we were specifically interested in examining the feedback systems of digital platforms as a means to capture public sentiment. Analysis of YouTube comments on TED talks and their (dis-) likes showed that while public sentiment is positive for non-white speakers, it is negative for women and talks about gender and ethnicity related inequalities.

Regarding descriptive representation, it seems as if TED media increased their efforts to achieve a more balanced gender representation in TED talks. The share of female speakers was below one third in 2006, but in the first half of 2017, more woman than men gave TED talks. One possible reason is that this was a reaction to an earlier study about scientific careers and TED talks (Sugimoto, Thelwall, et al. 2013), which was covered in the media (A. Taylor 2013) and revealed an overall low share of female speakers. Nevertheless, while the representation of women shows improvements, the situation is different for non-white ethnic groups in TED talks. Only one in five TED talks is given by a non-white speaker with no evidence for improvement over time. Digital content providers like TED media should increase their efforts to prevent that talking about science and important matters of societal change on a global stage remains a privilege of white people. Otherwise, under-representation of certain ethnic groups in the digital sphere can, similar to traditional media sources, further enhance stereotypes and negative attitudes.

When it comes to public sentiment, we showed that female presenters receive more negative feedback on YouTube than male speakers. To our surprise, the sentiment of non-white speakers was on average more positive than for white speakers and this finding holds for additional robustness checks. There is no reason to assume that non-white speakers in general perform better at stage and negative attitudes could rather be expected for other ethnic groups and not for white speakers. Future research could enhance our knowledge about this puzzling finding.

Regarding substantive representation, our findings provide evidence for a negative public sentiment of talks about inequalities related to women and certain ethnic groups. These talks often contain depressing rather than entertaining content, which might to some extent be reflected in video dislikes and comment sentiments. Some of the YouTube comments in our sample, as hinted by one of our examples above, contain very harsh language. Nevertheless, they are visible to all users, which raises concerns about YouTube's decision making related to removing obnoxious content.

Although computational methods enable promising social science research as demonstrated in this paper, their use also comes with important limitations. Regarding the image recognition algorithm that we used to measure descriptive representation, it is only possible to retrieve binary annotations for the sex of persons. These simplified annotations may not be in line with the gender a speaker identifies with. Related to this, gender is increasingly considered as a non-binary social construct. Likewise, as we mentioned above, the algorithm is limited to recognizing only certain ethnic groups and is unable to fully grasp the concept of ethnicity, which is also a complex social construct. Furthermore, the share of non-white speakers in TED talks was so low that we had to aggregate all non-white categories. With regards to the measurement of public sentiment on platforms like YouTube, our knowledge of the user population is very limited. We cannot assess whether only specific people on YouTube, e.g. people with negative attitudes towards white speakers, prefer to engage via commenting or liking. Nevertheless, regardless of how one interacts publicly on YouTube, these interactions are always visible on the platform. For this reason, quantification of the public sentiment in the digital sphere is still useful, even though we do not know what kind of users produce this content. Nevertheless we are only able to measure sentiment of users that actually commented on talks. Future studies could rely on experimental research designs to investigate the sentiment of *silent users* and how commenting behavior can impact their opinion.

With regards to the use of image recognition systems, researchers should also consider ethical implications. The authors of this paper would like to highlight that we agree with many citizens, scholars and politicians who consider some applications, e.g. for mass surveillance systems, troublesome and unethical. We fully support initiatives like *Safe Face Pledge*, which provides guidelines for ethic principles of facial analysis technology. However, we still think it is worth to study whether image recognition can also be utilized for good causes, which is why we examined whether its performance is reliable enough for social science research.

Despite the limitations of our work, this paper contributes to the emerging body of literature about the representation of disadvantaged groups in the digital sphere. As a substantial part thereof, TED talks are a central outlet for popularizing science and they reach millions of people around the globe. Our results raise some concerns, particularly about the representation of certain ethnic groups in these talks. This highlights the importance of speaker diversity to reduce stereotypes about scientists and people driving societal change. However, our knowledge about new features of these platforms, like interactive feedback systems and their utilization for digitally interacting with content by or about minorities, is still limited. We encourage scholars to further examine such feedback systems and the public sentiment of women end ethnicity representation on digital platforms.

References

Anderson, Ashley A., Dominique Brossard, Dietram A. Scheufele, Michael A. Xenos, and Peter Ladwig (2014). "The "Nasty Effect:" Online Incivility and Risk Perceptions of Emerging Technologies". In: *Journal of Computer-Mediated Communication* 19.3, pp. 373–387. ISSN: 10836101. DOI: 10.1111/jcc4.12009.

Bar-Gill, Sagit and Neil Gandal (2017). "Online Exploration, Content Choice& Echo Chambers: An Experiment". In: *CEPR Discussion Paper No. DP11909.* URL: https://ssrn.com/abstract=2937758.

Baum, Matthew A. (2004). "Circling the Wagons: Soft News and Isolationism in American Public Opinion". In: *International Studies Quarterly* 48.2, pp. 313–338.

Benkler, Yochai (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* New Haven: Yale University Press. ISBN: 9780300125771.

Benoit, Kenneth (2018). *quanteda: Quantitative Analysis of Textual Data.* DOI: 10.5281/zenodo.1004683. URL: http://github.com/kbenoit/quanteda.

Bleich, Erik, Irene Bloemraad, and Els de Graauw (2015). "Migrants, Minorities and the Media: Information, Representations and Participation in the Public Sphere". In: *Journal of Ethnic and Migration Studies* 41.6, pp. 857–873. ISSN: 14699451. DOI: 10.1080/1369183X.2014.1002197.

Boukes, Mark and Hajo G. Boomgaarden (2015). "Soft News With Hard Consequences? Introducing a Nuanced Measure of Soft Versus Hard News Exposure and Its Relationship With Political Cynicism". In: *Communication Research* 42.5, pp. 701–731. ISSN: 0093-6502. DOI: 10.1177/0093650214537520.

Brackeen, Brian (2018). *We've Retired Our Diversity Recognition App: Here's Why.* Kairos Blog. URL: https://www.kairos.com/blog/we-ve-retired-our-diversity-recognition-app-here-s-why.

Buchanan, Tom (2015). "Aggressive priming online: Facebook adverts can prime aggressive cognitions". In: *Computers in Human Behavior* 48, pp. 323–330. ISSN: 07475632.

Buolamwini, Joy and Timnit Gebru (2018). "Gender shades: Intersectional accuracy disparities in commercial gender classification". In: *Conference on Fairness, Accountability and Transparency*, pp. 77–91.

Celis, Karen and Sarah Childs (2008). "Introduction: The Descriptive and Substantive Representation of Women: New Directions". In: *Parliamentary Affairs* 61.3, pp. 419–425. ISSN: 0031-2290. DOI: 10.1093/pa/gsn006.

Celis, Karen and Silvia Erzeel (2015). "Beyond the Usual Suspects: Non-Left, Male and Non-Feminist MPs and the Substantive Representation of Women". In: *Government and Opposition* 50.1, pp. 45–64. DOI: 10.1017/gov.2013.42.

Chambers, David Wade (1983). "Stereotypic images of the scientist: The draw-a-scientist test". In: *Science Education* 67.2, pp. 255–265. ISSN: 0036-8326. DOI: 10.1002/sce.3730670213. URL: http://doi.wiley.com/10.1002/sce.3730670213.

Chetty, Naganna and Sreejith Alathur (2018). "Hate speech review in the context of online social networks". In: *Aggression and Violent Behavior* 40, pp. 108–118. DOI: 10.1016/j.avb.2018.05.003.

Dahl, Robert A. (1971). *Polyarchy: Participation and Opposition.* New Haven: Yale University. ISBN: 0300015658.

Dearden, Lizzie (2018). "Facial recognition to be deployed by police across London, sparking human rights concerns". In: URL: https://www.independent.co.uk/news/uk/home-news/facial-recognition-london-police-accuracy-human-rights-crime-database-a8422056.html.

Dominus, Susan (2017). *When the Revolution Came for Amy Cuddy.* URL: https://www.nytimes.com/2017/10/18/magazine/when-the-revolution-came-for-amy-cuddy.html.

Dovi, Suzanne (2002). "Preferable descriptive representatives: Will just any woman, black, or latino do?" In: *American Political Science Review* 96.4, pp. 729–743. ISSN: 15375943. DOI: 10.1017/S0003055402000412.

Entman, Robert M. (1993). "Framing: Toward Clarification of a Fractured Paradigm". In: *Journal of Communication* 43.4, pp. 51–58. ISSN: 00219916.

Entman, Robert M. and Nikki Usher (2018). “Framing in a Fractured Democracy: Impacts of Digital Technology on Ideology, Power and Cascading Network Activation”. In: *Journal of Communication* 68.2, pp. 298–308. ISSN: 00219916. DOI: 10.1093/joc/jqx019.

Feezell, Jessica T. (2018). “Agenda Setting through Social Media: The Importance of Incidental News Exposure and Social Filtering in the Digital Era”. In: *Political Research Quarterly* 71.2, pp. 482–494. DOI: 10.1177/1065912917744895.

Fonda, Jane, Robin Morgan, and Gloria Steinem (2017). *The Status of Women in the U.S. Media 2017*. Tech. rep. Women’s Media Center. URL: https://www.womensmediacenter.com/reports/the-status-of-women-in-u.s.-media-2017.

Gelman, Andrew (2008). “Scaling regression inputs by dividing by two standard deviations”. In: *Statistics in Medicine* 27.15, pp. 2865–2873. ISSN: 02776715. DOI: 10.1002/sim.3107. URL: http://doi.wiley.com/10.1002/sim.3107.

Giussani, Bruno (2015). *Even the tech savvy should turn off their phones and listen*. URL: https://www.theguardian.com/commentisfree/2015/jun/21/ted-with-no-twitter-no-smartphones-experiment.

Grimmer, Justin and Brandon M Stewart (2012). “Text as Data: The Promise and Pitfals of Automatic Content Analysis Methods for Political Texts”. In: *Political Analysis* 21.617, pp. 267–297. ISSN: 1047-1987, 1476-4989. DOI: 10.1093/pan/mps028. URL: http://www.stanford.edu/~jgrimmer/tad2.pdf.

Gunther, Albert C. and Stella Chih-Yun Chia (2001). “Predicting Pluralistic Ignorance: The Hostile Media Perception and Its Consequences”. In: *Journalism & Mass Communication Quarterly* 78.4, pp. 688–701. ISSN: 1077-6990.

Guo, Lei and Summer Harlow (2014). “User-Generated Racism: An Analysis of Stereotypes of African Americans, Latinos, and Asians in YouTube Videos”. In: *Howard Journal of Communications* 25.3, pp. 281–302. ISSN: 10964649. DOI: 10.1080/10646175.2014.925413.

Hargittai, Eszter (2010). "Digital Na(t)ives? Variation in Internet Skills and Uses among Members of the "Net Generation"". In: *Sociological Inquiry* 80.1, pp. 92–113. ISSN: 0038-0245. DOI: 10.1111/j.1475-682X.2009.00317.x.

Hargittai, Eszter, Lindsay Fullerton, Ericka Menchen-Trevino, and Kristin Yates Thomas (2010). "Trust Online: Young Adults' Evaluation of Web Content". In: *International Journal of Communication* 4, pp. 468–494.

Harp, Dustin, Ingrid Bachmann, and Jaime Loke (2014). "Where Are the Women? The Presence of Female Columnists in U.S. Opinion Pages". In: *Journalism & Mass Communication Quarterly* 91.2, pp. 289–307. ISSN: 1077-6990. DOI: 10.1177/1077699014527457. URL: http://journals.sagepub.com/doi/10.1177/1077699014527457.

Hindman, Matthew (2008). "What Is the Online Public Sphere Good For?" In: *The Hyperlinked Society*. Ed. by Joseph Turow and Lokman Tsui. Ann Arbor: University of Michigan Press, pp. 268–288. ISBN: 978-0-472-07043-5.

— (2009). *The Myth of Digital Democracy*. Princeton: Princeton University Press. ISBN: 978-0-691-13868-8.

— (2018). *The Internet Trap: How the Digital Economy Builds Monopolies and Undermines Democracy*. Princeton: Princeton University Press. ISBN: 978-0-691-15926-3.

Hu, Minqing and Bing Liu (2004). "Mining Opinion Features in Customer Reviews". In: *19th national conference on Artifical intelligence*, pp. 755–760. ISSN: 1581138889. DOI: 10.1145/1014052.1014073.

Huberman, Bernardo A., Peter L. T. Pirolli, James E. Pitkow, and Rajan M. Lukose (1998). "Strong Regularities in World Wide Web Surfing". In: *Science* 280, pp. 95–97. ISSN: 0036-8075.

Imai, Kosuke and Kabir Khanna (2016). "Improving ecological inference by predicting individual ethnicity from voter registration records". In: *Political Analysis* 24.2, pp. 263–272. DOI: 10.1093/pan/mpw001.

Iyengar, Shanto and Donald R. Kinder (1987). *News That Matters: Television and American Opinion*. Chicago: University of Chicago Press.

Jia, Sen, Thomas Lansdall-Welfare, Saatviga Sudhahar, Cynthia Carter, and Nello Cristianini (2016). “Women are seen more than heard in online newspapers”. In: *PLoS ONE* 11.2. Ed. by Keith Laws, e0148434. ISSN: 19326203. DOI: 10.1371/journal.pone.0148434.

Jockers, Matthew L. (2017). *Syuzhet: Extract sentiment and plot arcs from Text.* URL: https://github.com/mjockers/syuzhet.

Johnson, Eric J., Steven Bellman, and Gerald L. Lohse (2003). “Cognitive Lock-In and the Power Law of Practice”. In: *Journal of Marketing* 67.2, pp. 62–75. ISSN: 00222429.

Kedrowicz, April A. and Julie L. Taylor (2016). “Shifting rhetorical norms and electronic eloquence: TED talks as formal presentations”. In: *Journal of Business and Technical Communication* 30.3, pp. 352–377. ISSN: 15524574. DOI: 10.1177/1050651916636373.

Knox, Dean and Christopher Lucas (2018). “A Dynamic Model of Speech for the Social Sciences”. URL: https://asiapolmeth.princeton.edu/sites/default/files/polmeth/files/deanknox.pdf.

Ksiazek, Thomas B., Limor Peer, and Kevin Lessard (2016). “User engagement with online news: Conceptualizing interactivity and exploring the relationship between online news videos and user comments”. In: *New Media & Society* 18.3, pp. 502–520. ISSN: 1461-4448. DOI: 10.1177/1461444814545073.

Liu, Bing (2012). “Sentiment Analysis and Opinion Mining”. In: *Synthesis Lectures on Human Language Technologies* 5.1, pp. 1–167. ISSN: 1947-4040. DOI: 10.2200/S00416ED1V01Y201204HLT016. URL: http://www.morganclaypool.com/doi/abs/10.2200/S00416ED1V01Y201204HLT016.

Lucas, Christopher, Richard A Nielsen, Margaret E Roberts, Brandon M Stewart, Alex Storer, and Dustin Tingley (2015). “Computer-Assisted Text Analysis for Comparative Politics”. In: *Political Analysis* 23.2, pp. 254–277. ISSN: 1047-1987. DOI: 10.1093/pan/mpu019. URL: http://pan.oxfordjournals.org/cgi/doi/10.1093/pan/mpu019%5Cnpapers3://publication/doi/10.1093/pan/mpu019.

Lüdecke, Daniel (2018). *ggeffects: Create Tidy Data Frames of Marginal Effects for 'ggplot' from Model Outputs.* URL: https://github.com/strengejacke/ggeffects.

Ludewig, Julia (2017). "TED Talks as an Emergent Genre". In: *CLCWeb: Comparative Literature and Culture* 19.1.

MacDonald, Jason A. and Erin E. O'Brien (2011). "Quasi-Experimental Design, Constituency, and Advancing Women's Interests: Reexamining the Influence of Gender on Substantive Representation". In: *Political Research Quarterly* 64.2, pp. 472–486. DOI: 10.1177/1065912909354703.

Mansbridge, Jane (1999). "Should Blacks Represent Blacks and Women Represent Women? A Contingent "Yes"". In: *The Journal of Politics* 61.03, p. 628. ISSN: 0022-3816. DOI: 10.2307/2647821.

Martins, Nicole and Kristen Harrison (2012). "Racial and Gender Differences in the Relationship Between Children's Television Use and Self-Esteem". In: *Communication Research* 39.3, pp. 338–357. ISSN: 0093-6502. DOI: 10.1177/0093650211401376. URL: http://journals.sagepub.com/doi/10.1177/0093650211401376.

Matamoros-Fernández, Ariadna (2017). "Platformed racism: the mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube". In: *Information Communication and Society* 20.6, pp. 930–946. ISSN: 14684462. DOI: 10.1080/1369118X.2017.1293130. URL: https://www.tandfonline.com/doi/abs/10.1080/1369118X.2017.1293130?journalCode=rics20.

McCombs, Maxwell and Amy Reynolds (2009). "How the News Shapes Our Civic Agenda". In: *Media Effects*. Ed. by Jennings Bryant and Mary Beth Oliver. New York: Routledge, pp. 1–16. ISBN: 978-0-8058-6449-6.

Merskin, Debra (2017). "Media Representation: Minorities". In: *The International Encyclopedia of Media Effects*. Ed. by Patrick Rössler, Cynthia A. Hoffner, and Liesbet van Zoonen. Chichester, West Sussex and Malden, MA: John Wiley & Sons Inc., pp. 1096–1106. ISBN: 9781118784044. DOI: 10.1002/9781118783764.wbieme0147.

Miller, David I., Kyle M. Nolla, Alice H. Eagly, and David H. Uttal (2018). "The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies". In: *Child Development*. ISSN: 00093920. DOI: 10.1111/cdev.13039. URL: http://doi.wiley.com/10.1111/cdev.13039.

Miller, John (2006). “Who’s Telling the news: racial representation among news Gatherers in Canada’s Daily newsrooms”. In: *international Journal of Diversity in Organizations, Communities and nations* 4, pp. 133–142.

Murray, Kyle B. and Gerald Häubl (2007). “Explaining Cognitive Lock-In: The Role of Skill-Based Habits of Use in Consumer Choice”. In: *Journal of Consumer Research* 34.1, pp. 77–88. ISSN: 0093-5301.

Peer, Limor and Thomas B. Ksiazek (2011). “YouTube and the Challenge to Journalism: New standards for news videos online”. In: *Journalism Studies* 12.1, pp. 45–63. ISSN: 1461-670X. DOI: 10.1080/1461670X.2010.511951.

Pitkin, Hanna Fenichel (1967). *The Concept of Representation.* Berkeley: University of California Press. ISBN: 0520021568.

Pond, Philip and Jeff Lewis (2019). “Riots and Twitter: connective politics, social media and framing discourses in the digital public sphere”. In: *Information, Communication & Society* 22.2, pp. 213–231. ISSN: 1369-118X. DOI: 10.1080/1369118X.2017.1366539.

Puschmann, Cornelius and Alison Powell (2018). “Turning Words Into Consumer Preferences: How Sentiment Analysis Is Framed in Research and the News Media”. In: *Social Media + Society* 4.3, p. 205630511879772. ISSN: 2056-3051. DOI: 10.1177/2056305118797724. URL: http://journals.sagepub.com/doi/10.1177/2056305118797724.

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* Vienna, Austria. DOI: 10.1038/sj.hdy.6800737. URL: https://www.r-project.org/.

Reynolds, Glenn H. (2006). *An Army of Davids: How Markets and Technology Empower Ordinary People to Beat Big Media, Big Government, and Other Goliaths.* Nashville: Thomas Nelson. ISBN: 1-595-55054-2.

Rieder, Bernhard (2015). *YouTube Data Tools.* URL: https://tools.digitalmethods.net/netvizz/youtube/index.php.

Rinker, Tyler (2018). *sentimentr: Calculate Text Polarity Sentiment.* DOI: 10.5281/zenodo.222103. URL: http://github.com/trinker/sentimentr.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-luis, Shana Kushner Gadarian, Bethany Albertson, and David G Rand (2014). "Structural Topic Models for Open-Ended Survey Responses". In: *American Journal of Political Science* 58.4, pp. 1064–1082. DOI: 10.1111/ajps.12103.

Rothschild, David and Neil Malhotra (2014). "Are public opinion polls self-fulfilling prophecies?" In: *Research & Politics* 1.2, pp. 1–10. ISSN: 2053-1680. DOI: 10.1177/2053168014547667.

Scheufele, Dietram (2018). "(New) political interfaces in the life sciences". In: *Politics and the Life Sciences* 37.1, pp. 78–87. DOI: 10.1017/pls.2017.9.

Scheufele, Dietram A. (1999). "Framing as a Theory of Media Effects". In: *Journal of Communication* 49.1, pp. 103–122. ISSN: 00219916.

Scheufele, Dietram A. and David Tewksbury (2007). "Framing, Agenda Setting, and Priming: The Evolution of Three Media Effects Models". In: *Journal of Communication* 57.1, pp. 9–20. ISSN: 00219916.

Schradie, Jen (2018). "The Digital Activism Gap: How Class and Costs Shape Online Collective Action". In: *Social Problems* 65.1, pp. 51–74. ISSN: 0037-7791. DOI: 10.1093/socpro/spx042. URL: https://academic.oup.com/socpro/article/65/1/51/4795348.

Schwartz, Ariel (2018). *What it's like to attend the TED talks, where attendees pay $10,000 to learn the next big ideas.* URL: https://www.businessinsider.de/ted-2017-conference-photos-2017-4.

Schwemmer, Carsten (2018). *stminsights. A 'Shiny' Application for Inspecting Structural Topic Models.* URL: https://cran.r-project.org/web/packages/stminsights/index.html.

Shor, Eran, Arnout van de Rijt, Alex Miltsov, Vivek Kulkarni, and Steven Skiena (2015). "A Paper Ceiling: Explaining the Persistent Underrepresentation of Women in Printed News". In: *American Sociological Review* 80.5, pp. 960–984. ISSN: 19398271. DOI: 10.1177/0003122415596999.

Singh Chawla, Dalmeet (2016). “Textual analysis of science-related TED Talks”. In: *The Winnower* January 29, 2016. URL: https://thewinnower.com/papers/3338-textual-analysis-of-science-related-ted-talks.

Sobieraj, Sarah (2017). “Bitch, slut, skank, cunt: patterned resistance to women’s visibility in digital publics”. In: *Information, Communication & Society*, pp. 1–15. ISSN: 1369-118X. DOI: 10.1080/1369118X.2017.1348535. URL: https://www.tandfonline.com/doi/full/10.1080/1369118X.2017.1348535.

Soroka, Stuart N. and Christopher Wlezien (2010). *Degrees of Democracy: Politics, Public Opinion, and Policy.* Cambridge: Cambridge University Press.

Sugimoto, Cassidy R. and Mike Thelwall (2013). “Scholars on Soap Boxes: Science Communication and Dissemination in TED Videos”. In: *Journal of the American Society for Information Science and Technology* 64.4, pp. 663–674. ISSN: 14923831. DOI: 10.1002/asi.22764.

Sugimoto, Cassidy R., Mike Thelwall, Vincent Larivière, Andrew Tsou, Philippe Mongeon, and Benoit Macaluso (2013). “Scientists Popularizing Science: Characteristics and Impact of TED Talk Presenters”. In: *PLoS ONE* 8.4. Ed. by Judit Bar-Ilan, e62403. ISSN: 19326203. DOI: 10.1371/journal.pone.0062403.

Sundar, S. Shyam, Haiyan Jia, T. Franklin Waddell, and Yan Huang (2015). “Toward a Theory of Interactive Media Effects (TIME): Four Models for Explaining How Interface Features Affect User Psychology”. In: *The Handbook of the Psychology of Communication Technology.* Ed. by S. Shyam Sundar. Chichester: Wiley Blackwell, pp. 47–86. ISBN: 9781118426524.

Taylor, Alison (2013). *TED talks rising to the challenge of gender imbalance.* URL: https://www.piquenewsmagazine.com/whistler/ted-talks-rising-to-the-challenge-of-gender-imbalancerecent-study-shines-spotlight-on-the-high-ratio-of-men-speaking-compared-to-small-number/Content?oid=2467758.

Ter Wal, Jessika, ed. (2002). *Racism and Cultural Diversity in the Mass Media: An overview of research and examples of good practice in the EU Member States, 1995-2000*. Vienna: European Monitoring Centre on Racism and Xenophobia (EUMC).

Tewksbury, David and Dietram A. Scheufele (2009). "News Framing Theory and Research". In: *Media Effects*. Ed. by Jennings Bryant and Mary Beth Oliver. New York: Routledge, pp. 17–33. ISBN: 978-0-8058-6449-6.

Tsou, Andrew, Mike Thelwall, Philippe Mongeon, and Cassidy R. Sugimoto (2014). "A Community of Curious Souls: An Analysis of Commenting Behavior on TED Talks Videos". In: *PLoS ONE* 9.4, e93609. ISSN: 19326203. DOI: 10.1371/journal.pone.0093609.

Turnaround Management Association (2017). *Attendee List. Ted Talks Time Two. Bank Street, Stamford, CT. April 24, 2017*. URL: https://turnaround.org/sites/default/files/AttendeeList-TedTalks-4.24.17.pdf.

U.S. Census Bureau (2019). *2013-2017 American Community Survey 5-Year Estimates*.

Valenzuela, Sebastián, Martina Piña, and Josefina Ramírez (2017). "Behavioral Effects of Framing on Social Media Users: How Conflict, Economic, Human Interest, and Morality Frames Drive News Sharing". In: *Journal of Communication* 67.5, pp. 803–826. ISSN: 00219916. DOI: 10.1111/jcom.12325.

Valkenburg, Patti M., Jochen Peter, and Joseph B. Walther (2016). "Media Effects: Theory and Research". In: *Annual review of psychology* 67, pp. 315–338. DOI: 10.1146/annurev-psych-122414-033608.

van Duyn, Emily and Jessica Collier (2019). "Priming and Fake News: The Effects of Elite Discourse on Evaluations of News Media". In: *Mass Communication and Society* 22.1, pp. 29–48. ISSN: 1520-5436. DOI: 10.1080/15205436.2018.1511807.

Van Klingeren, Marijn, Hajo G. Boomgaarden, Rens Vliegenthart, and Claes H. De Vreese (2015). "Real world is not enough: The media as an additional source of negative attitudes toward immigration, Comparing Denmark and the Netherlands". In: *European Sociological Review* 31.3, pp. 268–283. ISSN: 14682672. DOI: 10.1093/esr/jcu089.

Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady (1995). *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge: Harvard University Press. ISBN: 0674942930.

Wais, Kamil (2016). "Gender Prediction Methods Based on First Names with genderizeR". In: *The R Journal* 8.1, pp. 17–37. DOI: 10.32614/RJ-2016-002. URL: https://doi.org/10.32614/RJ-2016-002.

Wickham, Hadley (2016). *tidyverse: Easily Install and Load 'Tidyverse' Packages*. URL: https://cran.r-project.org/package=tidyverse.

Williams, Christopher J. and Martijn Schoonvelde (2018). "It Takes Three: How Mass Media Coverage Conditions Public Responsiveness to Policy Outputs in the United States". In: *Social Science Quarterly* 99.5, pp. 1627–1636. ISSN: 00384941. DOI: 10.1111/ssqu.12525.

Wlezien, Christopher (1995). "The Public as Thermostat: Dynamics of Preferences for Spending". In: *American Journal of Political Science* 39.4, pp. 981–1000. ISSN: 00925853. DOI: 10.2307/2111666.

Xenos, Michael A., Timothy Macafee, and Antoinette Pole (2017). "Understanding variations in user response to social media campaigns: A study of Facebook posts in the 2010 US elections". In: *New Media & Society* 19.6, pp. 826–842. ISSN: 1461-4448. DOI: 10.1177/1461444815616617.

Zaller, John (2003). "A New Standard of News Quality: Burglar Alarms for the Monitorial Citizen". In: *Political Communication* 20.2, pp. 109–130. ISSN: 1058-4609.

Whose ideas are worth spreading? The representation of women and ethnic groups in TED talks - Supporting Information

Replication material for this article is available at Harvard Dataverse `https://doi.org/10.7910/DVN/EUDWP3`.

S1 Image recognition algorithm

This section contains metrics on the validation of the image recognition algorithm that we used to capture sex and ethnicity of TED speakers. Numbers are based on a comparison of human and algorithmic coding of 200 random samples speaker images.

Validation

The following tables show the pairwise inter-rater agreements between human coders and the algorithm.

Table S1: Pairwise inter-rater agreement: ethnicity

Rater 1	Rater 2	Fleiss' Kappa
Human coder 1	Algorithm	0.833
Human coder 2	Algorithm	0.888
Human coder 1	Human coder 2	0.825

The average inter-rater agreement for ethnicity between human coders and the image recognition algorithm is 0.860.

Table S1: Pairwise inter-rater agreement: gender

Rater 1	Rater 2	Fleiss' Kappa
Human 1	Algorithm	0.946
Human 2	Algorithm	0.968
Human 1	Human 2	0.957

The average inter-rater agreement for sex between human coders and the image recognition algorithm is 0.957.

The following tables show overall inter-rater agreements between both human coders and the algorithm by ethnicity and sex.

Table S1: Overall inter-rater agreements by label: ethnicity

Label	Fleiss' Kappa
Asian	0.932
Black	0.965
Hispanic	0.576
Other	0.707
White	0.864
Overall	0.848

Table S1: Overall inter-rater agreements by label: sex

Label	Fleiss' Kappa
Female	0.957
Male	0.957
Overall	0.957

In total, Kappa values for both ethnicity as well as sex are good, with high agreement for the majority of sample images about both attributes. The agreement for the ethnicity categories *Hispanic* and *Other* are worse in comparison to the remaining categories. It is worth highlighting at this point that the lower agreement can not only be ascribed to poor performance by the algorithm. Both, pairwise comparisons between human coders as well as between humans and the algorithm showed a small number of disagreements for these categories. For instance, some speaker images were labeled as *White* by one rater and *Hispanic* by another rater. Overall, our validation results show that annotations from the image recognition algorithm are not perfect, but mostly in line with annotations by human coders and sufficient for our research task.

Another approach for validating the algorithmic performance is to treat results from one human coder as a gold standard and assess the predictive performance of the image recognition algorithm. In doing so, we calculated precision, recall and F1 scores for each ethnicity class, as can be seen in the following table:

Table S1: Predictive performance by ethnicity class (image recognition algorithm)

Ethnicity Class	Precision	Recall	F1 Score
Asian	1.00	1.00	1.00
Black	0.95	0.95	0.95
Hispanic	0.54	0.46	0.50
Other	0.50	0.75	0.60
White	0.97	0.97	0.97

As for computing these metrics over all classes, several metrics could be used. The overall micro averaged F1 score, which does not take imbalanced class distributions into account, is at 0.96. In comparison, the macro averaged F1 score, which takes into account the sizes of all classes, is at 0.81.

For gender, precision is at 0.95, recall at 0.97 and the F1 score at 0.96.

Overall, these results are in line with our inter-rater metrics from above.

To further compare the performance of the image recognition algorithm with name-based approaches for identifying gender and ethnicity, we again treated human codings as a gold standard for ethnicity and gender. We then used the first names (for gender) and surnames (for ethnicity) to identify the most likely class for each category (see Wais 2016; Imai and Khanna 2016). Afterwards, we again computed precision, recall and F1 scores:

Table S1: Predictive performance by ethnicity class (name-based approach)

Ethnicity Class	Precision	Recall	F1 Score
Asian	0.81	0.47	0.60
Black	0.09	0.50	0.15
Hispanic	0.27	0.23	0.25
Other	0.00	0.00	0.00
White	0.90	0.83	0.87

The micro averaged F1 scores for the name-based predictions is at 0.86 and the macro averaged F1 score at 0.41. These results suggest that the image recognition algorithm performs substantially better for predicting ethnicity categories.
For the name-based gender predictions, precision, recall and F1 scores are at 0.94. These scores are also lower than those achieved by the image recognition algorithm, but only by a very small margin.

Probability distributions

The following figure shows the probability distributions for ethnicity annotations by the image recognition algorithm. Most of the probability mass is centered around 0% and 100%, which is why a conversion of the continuous measures to a binary non-white versus white indicator does not result in a major loss of information.

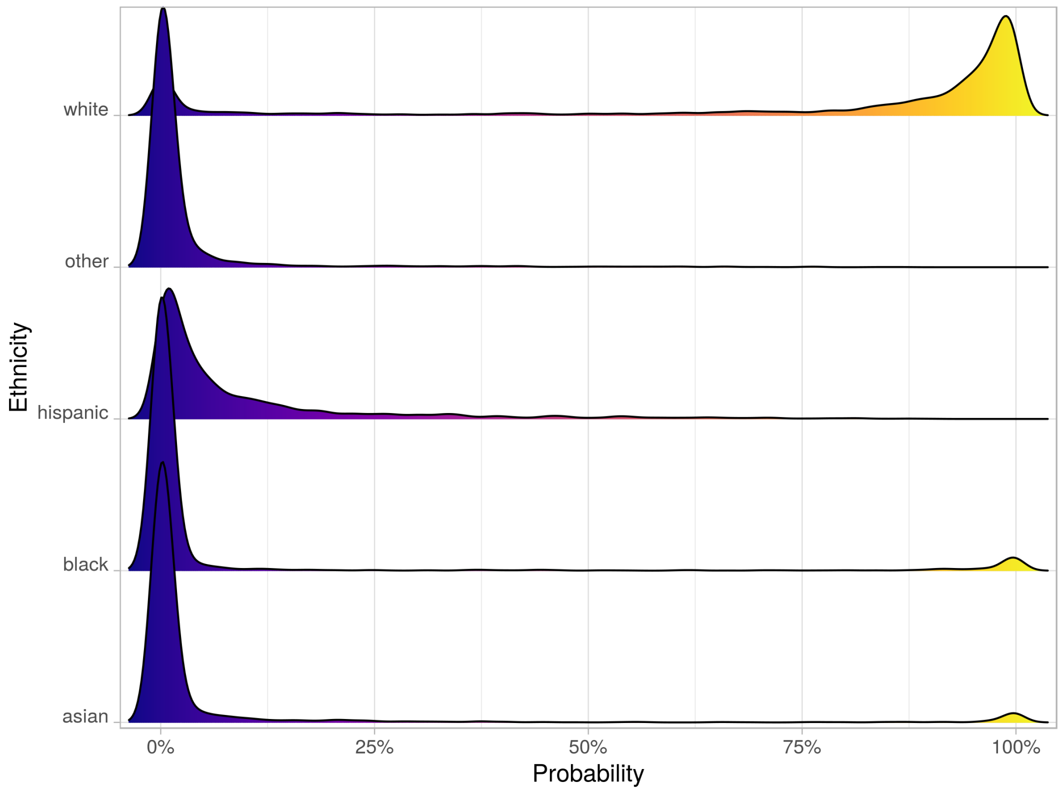

Figure S1: Probability distributions of image recognition algorithm by ethnicity

S2 Topic modeling

Topic model diagnostics

To find a model that best fits our research purpose, we computed three different models with 20, 30, and 50 topics. For each model, we then calculated semantic coherence and exclusivity, which are measures that quantify necessary statistical properties and are recommended by the authors of the structural topic model Roberts et al. 2014. Semantic coherence is higher when more probable words in a topic frequently co-occur together Mimno et al. 2011. Exclusivity is based on the FREX metric and achieves higher values when more words are exclusive to corresponding topics Lucas et al. 2015. The following figure illustrates average and median values for semantic coherence and exclusivity of all structural topic models that we fitted on the TED talk transcripts.

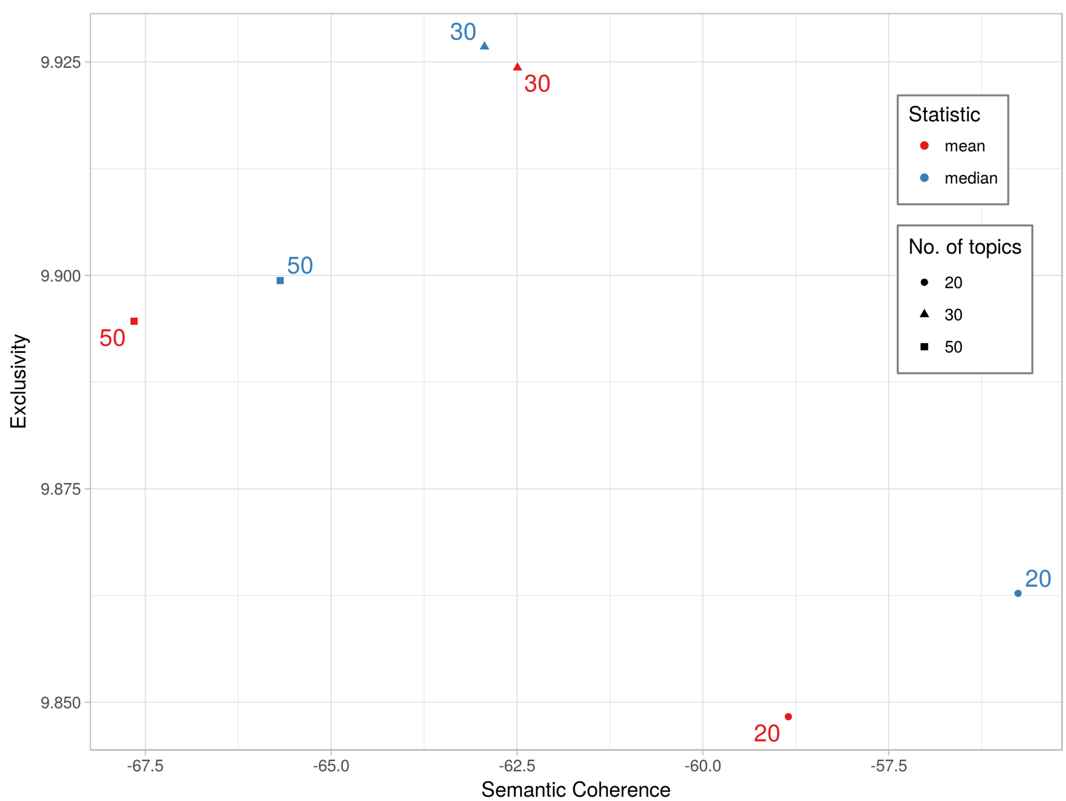

Figure S2: Probability distributions of image recognition algorithm by ethnicity

While the figure shows that no model is clearly superior, the model with 30 topics outperforms both other models in terms of exclusivity. In addition to statistical diagnostics, we inspected the topic models by analyzing the most frequent and exclusive words (Lucas et al. 2015, p. 19) - called *frex* terms - and representative talks with the highest proportions for each topic. Based upon frex terms and highly representative texts we at last assigned labels to each topic. During this evaluation procedure, the model with 30 topics also turned out to be the best model in relation to our research task and was therefore chosen as our final model.

Topic Labels and Proportions

Table S7: Topic labels, proportions and terms

Label	Proportion	FREX terms
family	8%	father, mother, famili, felt, mom, fear, love, told, knew, dream, met, son, home, sister, night, friend, dad, brother, die, cri
work & misc.	5%	guy, stuff, phone, somebodi, hey, everybodi, pick, box, laugh, email, anyway, morn, oh, sort, cartoon, room, shoe, fun, funni, check
computers & technology	5%	comput, machin, design, devic, technolog, softwar, interact, 3d, algorithm, digit, print, interfac, prototyp, manufactur, code, program, sensor, mit, screen, tool
countries & poverty	4%	china, africa, india, african, chines, countri, aid, incom, growth, poverti, econom, global, wealth, economi, gdp, inequ, europ, poor, asia, west
war & terror	4%	refuge, muslim, war, afghanistan, militari, peac, soldier, islam, conflict, weapon, arab, bomb, iraq, violenc, iran, terror, kill, gun, attack, secur
law & politics	4%	prison, vote, elect, crimin, polic, legal, citizen, law, democraci, polit, crime, justic, court, jail, govern, democrat, presid, corrupt, congress, lawyer
art	4%	artist, art, paint, museum, photograph, imag, draw, sculptur, color, visual, exhibit, photo, theater, beauti, portrait, photographi, studio, galleri, captur, wall
money & business	4%	market, busi, money, dollar, compani, financi, invest, sector, innov, profit, entrepreneur, buy, fund, bank, brand, product, employe, capit, custom, pay
general science	4%	choic, scienc, predict, decis, math, statist, scientist, wrong, knowledg, solv, answer, model, puzzl, evid, intellig, scientif, theori, bias, mathemat, choos
emotions & psychology	4%	compass, emot, self, happi, moral, psycholog, stress, empathi, relationship, autism, mental, sleep, desir, social, feel, suffer, smile, behavior, love, other
data & internet	4%	internet, onlin, twitter, web, data, media, network, googl, facebook, inform, privaci, digit, blog, post, content, websit, link, user, phone, site
medical	4%	cancer, patient, medic, clinic, treatment, doctor, tumor, vaccin, diseas, surgeri, medicin, drug, breast, health, physician, hospit, hiv, symptom, trial, epidem
games & music	3%	music, game, song, play, musician, sound, piano, player, sing, orchestra, listen, video, voic, instrument, opera, hear, nois, ear, improvis, concert
inequality	**3%**	**women, men, gender, gay, sexual, sex, girl, woman, rape, feminist, femal, black, marriag, male, boy, slaveri, marri, equal, abus, violenc**
universe & physics	3%	galaxi, particl, telescop, star, quantum, atom, univers, mathemat, hole, theori, dark, dimens, light, graviti, physicist, symmetri, physic, astronom, magnet, sun
astronautics	3%	mar, ice, earth, satellit, atmospher, asteroid, planet, cave, pole, moon, nasa, orbit, mountain, explor, glacier, rock, solar, mission, antarctica, surfac
genetics	3%	dna, gene, genom, cell, tissu, genet, molecul, stem, bacteria, protein, biolog, molecular, sequenc, chromosom, mutat, virus, silk, evolut, transplant, bone
architecture & design	3%	citi, architectur, urban, build, neighborhood, architect, park, street, mayor, built, hous, york, design, communiti, space, roof, site, town, tower, rio
environment	3%	oil, climat, energi, nuclear, carbon, fuel, emiss, coal, electr, gas, co2, solar, renew, wind, batteri, burn, fusion, effici, fossil, heat
language & writing	3%	languag, book, english, translat, write, text, word, dictionari, letter, spell, sentenc, read, librari, script, page, written, metaphor, writer, publish, editor
religion & mystery	3%	god, film, religion, movi, religi, fiction, storytel, charact, conscious, realiti, argument, stori, faith, magic, mysteri, truth, christian, tom, comic, believ
children & school	3%	school, teacher, educ, kid, student, children, classroom, teach, grade, colleg, class, parent, child, graduat, skill, adult, young, villag, lunch, childhood
animals & nature	3%	forest, chimpanze, tree, extinct, speci, eleph, creatur, bat, anim, ancestor, bird, amazon, rainforest, mammal, bear, monkey, soil, beetl, frog, ecosystem
body & sports	3%	leg, arm, finger, limb, disabl, breath, knee, foot, climb, danc, bodi, blind, athlet, swim, feet, wheelchair, ball, walk, injuri, jump
stopwords	2%	ca, ok, yeah, la, chris, ted, poem, flag, oh, mr, ah, card, anderson, yes, smell, prize, pleas, audienc, sir, prime
materials & sustainability	2%	plastic, water, sand, wast, oxygen, recycl, materi, bottl, mushroom, toilet, drink, air, wash, clean, chemic, pollut, sanit, river, temperatur, pump
food & farming	2%	bee, food, farmer, crop, mosquito, farm, eat, diet, bread, flower, seed, agricultur, meat, plant, feed, pig, meal, grain, chicken, cow
neuroscience & robotics	2%	neuron, brain, robot, ant, memori, cortex, neural, neurosci, activ, signal, pattern, region, task, motor, sensori, coloni, babi, function, rat, electrod
transportation	2%	car, vehicl, airplan, driver, crash, traffic, road, flight, drive, gps, fli, aircraft, balloon, seat, highway, mile, wheel, wing, pilot, plane
ocean & sea	2%	shark, fish, ocean, coral, whale, reef, sea, dolphin, boat, underwat, marin, ship, dive, island, shrimp, tag, pacif, swim, coast, jellyfish

Prevalence effects

The following figures includes prevalence effects from the structural topic model. Effects are illustrated for speaker ethnicity (A), speaker gender (B) and publication date of talks (C) with 95% confidence intervals.

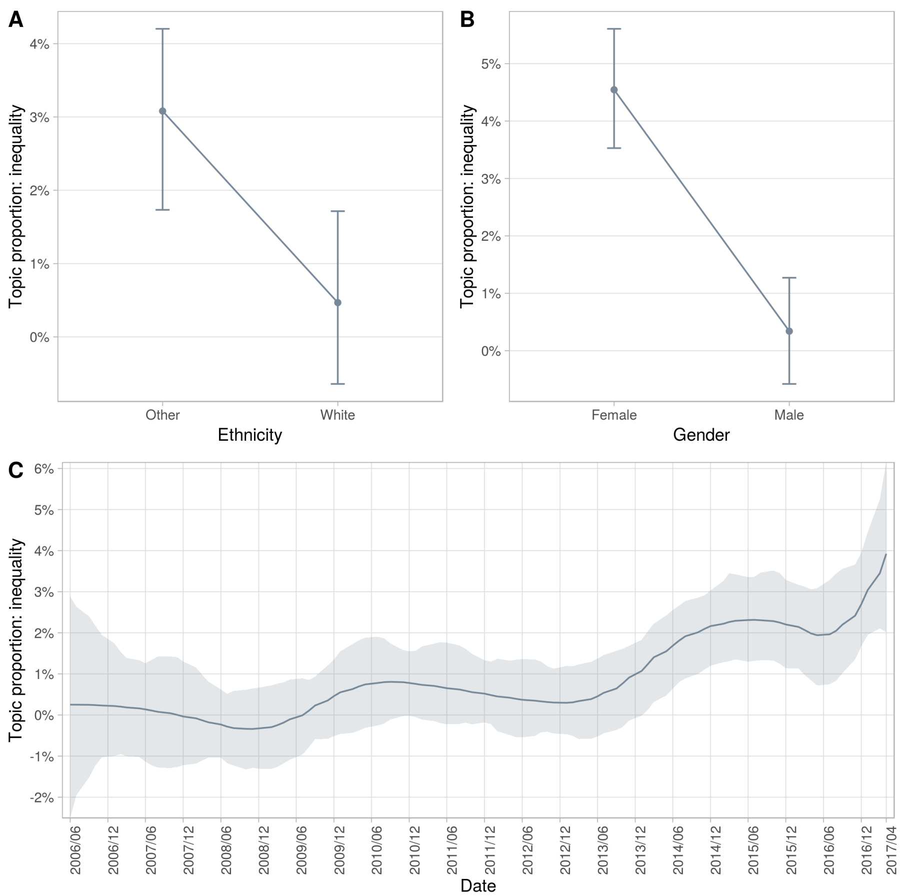

Figure S3: Prevalence effects of speaker ethnicity (A), speaker gender (B) and publication date of a talk (C) on inequality topic with 95% confidence intervals.

S3 Regressions

Coefficient table

The following table includes unstandardized regression coefficients for the comment sentiment and the number of dislikes of YouTube videos. The *work and misc.* topic was excluded from both models to avoid collinearity issues.

Table S8: Regression models for comment sentiment and number of dislikes

	Comment sentiment	Video dislikes
Intercept	$3.26\ (0.94)^{***}$	$-338.79\ (13.82)^{***}$
Video date	$-0.00\ (0.00)^{***}$	$0.17\ (0.01)^{***}$
Video views	$0.00\ (0.00)$	$0.00\ (0.00)^{***}$
Gender: female	$-0.01\ (0.00)^{**}$	$0.28\ (0.04)^{***}$
Ethnicity: non-white	$0.02\ (0.00)^{***}$	$-0.17\ (0.05)^{***}$
Topic: games and music	$0.03\ (0.02)$	$-0.02\ (0.31)$
Topic: stopwords	$-0.02\ (0.02)$	$-0.09\ (0.36)$
Topic: inequality	$-0.19\ (0.02)^{***}$	$3.53\ (0.34)^{***}$
Topic: countries and poverty	$-0.07\ (0.02)^{***}$	$0.39\ (0.28)$
Topic: universe and physics	$-0.03\ (0.02)$	$-0.18\ (0.28)$
Topic: family	$0.00\ (0.02)$	$-0.36\ (0.29)$
Topic: war and terror	$-0.16\ (0.02)^{***}$	$0.37\ (0.28)$
Topic: astronautics	$-0.06\ (0.02)^{**}$	$-1.05\ (0.30)^{***}$
Topic: genetics	$-0.04\ (0.02)$	$-0.91\ (0.30)^{**}$
Topic: architecture and design	$-0.01\ (0.02)$	$-1.16\ (0.30)^{***}$
Topic: environment	$-0.09\ (0.02)^{***}$	$0.42\ (0.30)$
Topic: materials and sustainability	$-0.02\ (0.02)$	$-1.33\ (0.34)^{***}$
Topic: law and politics	$-0.17\ (0.02)^{***}$	$1.01\ (0.30)^{***}$
Topic: art	$0.05\ (0.02)^{*}$	$-0.71\ (0.31)^{*}$
Topic: food and farming	$-0.05\ (0.02)^{*}$	$-0.09\ (0.33)$
Topic: language and writing	$0.01\ (0.02)$	$-0.30\ (0.35)$
Topic: religion and mystery	$-0.07\ (0.02)^{**}$	$2.39\ (0.36)^{***}$
Topic: money and business	$-0.03\ (0.02)$	$-0.87\ (0.31)^{**}$
Topic: neuroscience and robotics	$-0.04\ (0.02)$	$-1.31\ (0.31)^{***}$
Topic: children and school	$0.03\ (0.02)$	$-1.07\ (0.32)^{***}$
Topic: computers and technology	$-0.00\ (0.02)$	$-0.42\ (0.29)$
Topic: general science	$-0.02\ (0.02)$	$-0.15\ (0.31)$
Topic: emotions and psychology	$-0.01\ (0.02)$	$0.19\ (0.29)$
Topic: data and internet	$-0.05\ (0.02)^{*}$	$-0.39\ (0.30)$
Topic: transportation	$-0.09\ (0.02)^{***}$	$-0.53\ (0.35)$
Topic: animals and nature	$-0.02\ (0.02)$	$-1.43\ (0.31)^{***}$
Topic: body and sports	$-0.01\ (0.02)$	$-0.62\ (0.32)$
Topic: medical	$-0.09\ (0.02)^{***}$	$-1.50\ (0.28)^{***}$
Topic: ocean and sea	$-0.03\ (0.02)$	$-2.05\ (0.31)^{***}$
Log Likelihood	2913.20	−13709.98
Observations	2310	2323

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$; standard errors in parentheses.

Interaction between date and ethnicity

To further examine the relation between time and ethnicity regarding the YouTube sentiment, we computed an additional regression which includes an interaction between date and a dummy variable for white versus non-white speakers. The following visualization shows the corresponding effects while holding all other covariates at their observed values.

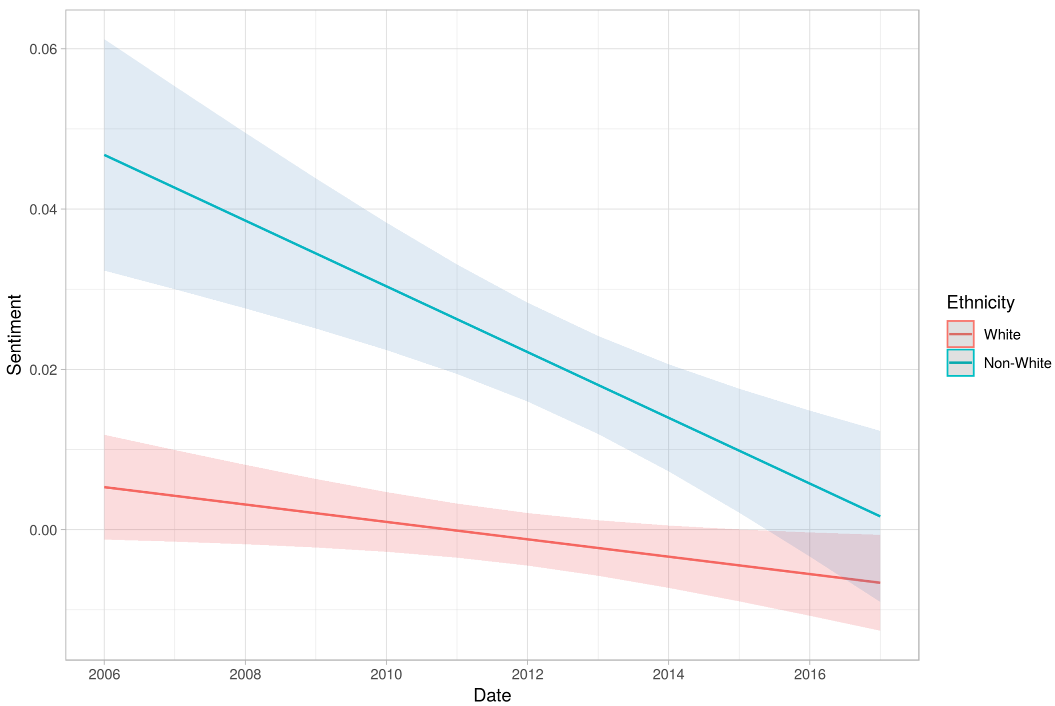

Figure S4: Predicted sentiment score for ethnicity and date.

It becomes apparent that, similar to our model without an interaction term, TED talks by non-white speakers receive a more positive sentiment on YouTube. However, over time, differences between non-white and white speakers vanish.

Dislikes on YouTube

The following figure includes regression estimates for the number of dislikes for each TED talk in our sample. Subfigure A shows standardized coefficients of generalized linear models with 95% confidence intervals. Subfigures B.1-B.3 show predicted values for descriptive and substantive representation covariates.

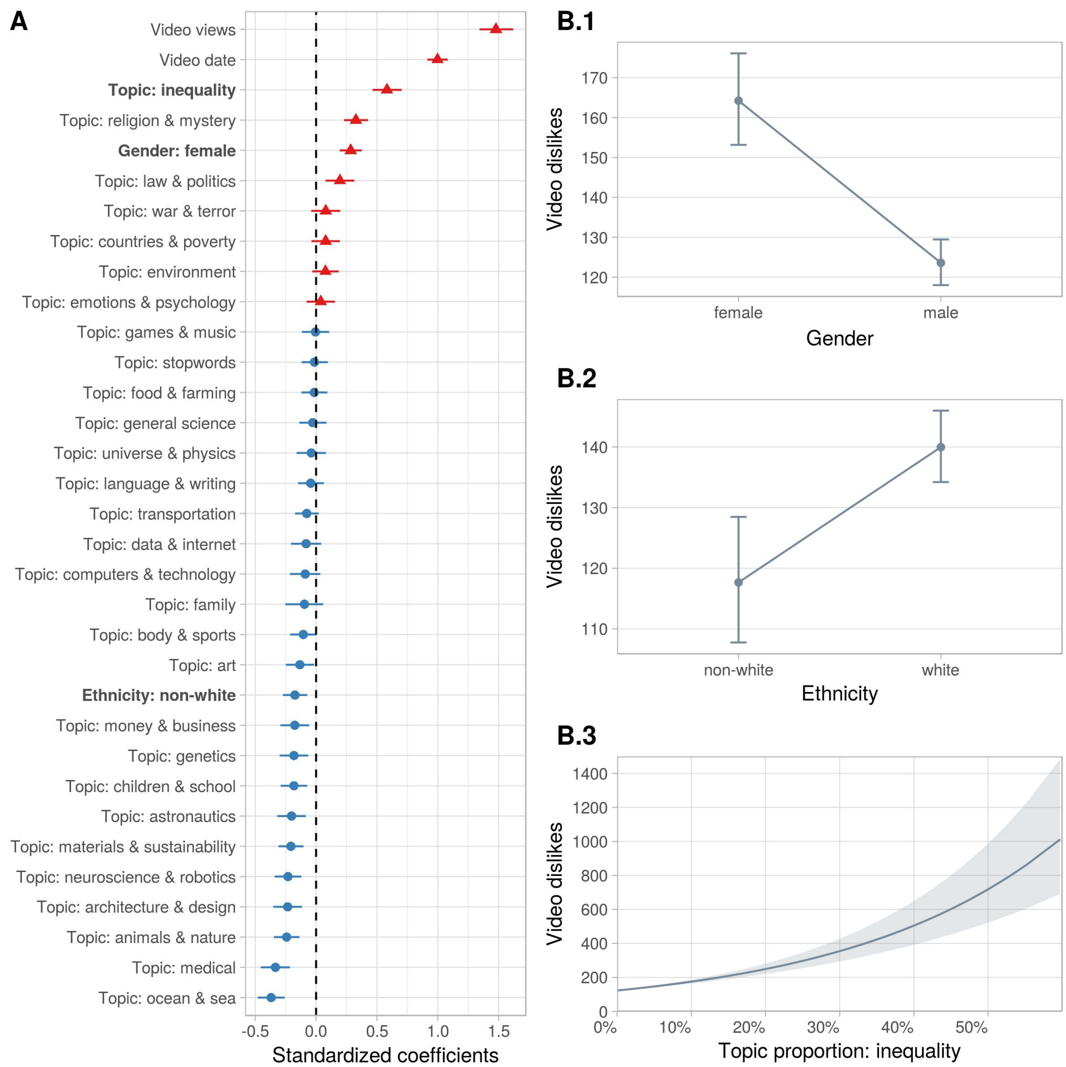

Figure S5: Regression estimates for the number of dislikes of TED YouTube videos. Sub-figure A: standardized coefficients of generalized linear model with 95% confidence intervals. Sub-figures B.1 - B.3: predicted sentiment values for descriptive and substantive representation covariates.

References

Imai, Kosuke and Kabir Khanna (2016). "Improving ecological inference by predicting individual ethnicity from voter registration records". In: *Political Analysis* 24.2, pp. 263–272. DOI: 10.1093/pan/mpw001.

Lucas, Christopher, Richard A Nielsen, Margaret E Roberts, Brandon M Stewart, Alex Storer, and Dustin Tingley (2015). "Computer-Assisted Text Analysis for Comparative Politics". In: *Political Analysis* 23.2, pp. 254–277. ISSN: 1047-1987. DOI: 10.1093/pan/mpu019. URL: http://pan.oxfordjournals.org/cgi/doi/10.1093/pan/mpu019%5Cnpapers3://publication/doi/10.1093/pan/mpu019.

Mimno, David, Hanna M. Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum (2011). "Optimizing semantic coherence in topic models". In: *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing.* 2, pp. 262–272. ISBN: 9781937284114.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-luis, Shana Kushner Gadarian, Bethany Albertson, and David G Rand (2014). "Structural Topic Models for Open-Ended Survey Responses". In: *American Journal of Political Science* 58.4, pp. 1064–1082. DOI: 10.1111/ajps.12103.

Wais, Kamil (2016). "Gender Prediction Methods Based on First Names with genderizeR". In: *The R Journal* 8.1, pp. 17–37. DOI: 10.32614/RJ-2016-002. URL: https://doi.org/10.32614/RJ-2016-002.