Do People Really Not Agree on What Can be Said?

Individual Differences in the Perception of Microaggressive, Derogatory and Hate Speech Against Women

Carsten Reinemann & Anna-Luisa Sacher

Abstract

In recent years, Western democracies have seen an increase in controversies about what can be said publicly. These controversies often lead to more general discussions about freedom of speech, "political correctness," "cancel culture," or the consequences of hateful and discriminatory speech spread on the Internet. However, till date, not much is known about the fundamental question of whether citizens perceive potentially harmful statements in similar ways and what might explain their varying perceptions. Against this backdrop, the current study investigates how microaggressive, derogatory, and hate speech against women are perceived depending on sociodemographics, experiences of discrimination, political attitudes, and trust in media. We develop a set of hypotheses and test these based on a standardized survey with a quota sample resembling the German population between 18 and 65 (N = 943). The survey included a split-ballot in which half of the respondents were asked to judge whether they regarded eight statements directed at women as acceptable and hurtful. The findings showed a great deal of consensus in the perception of those statements. While gender does not prove to be a key factor for explaining individual differences, age, experiences of discrimination, and media trust turn out to be significant predictors.

How free should speech be? How should free speech be? With these questions, Timothy Garton Ash (2016) outlines the tension between freedom of speech and discourse culture, which has become an issue of public debate in Western democracies in recent years. In the context of discussions about "political correctness" and "cancel culture," discriminatory and non-discriminatory speech, or the increase of hate speech on the Internet (e.g., Reimer, 2019), we have seen heated and controversial debates about whether certain statements fall within the space of what can be said and what

is socially or legally acceptable. However, the kind of statements being discussed, what exactly is criticized, and the standards used to judge varies throughout and is by no means completely obvious in every case: sometimes, it may be a matter of legal categories, i.e., the question of whether a statement constitutes a justiciable insult, incitement of the people, or even approval of a war of aggression; other discussions revolve around whether statements are discriminatory, misogynistic, anti-Semitic, Islamophobic, or disparaging in some other way; it may be a matter of whether statements represent disinformation or "fake news"; or the questions are discussed of whether statements should be disseminated (unchallenged) in the media or whether controversial content or accounts should be deleted from social networks.

On a more abstract level, these kinds of debates often turn into more general discussions about whether freedom of expression is increasingly in danger or whether, on the contrary, freedom of speech tends to be abused to a more frequent extent, especially in the online environment. From a jurisprudential perspective, these discussions are often referred to as the "democratic dilemma," which describes the tension between freedom of expression and freedom from discrimination and is considered one of the greatest challenges for modern democracy (e.g., Marker, 2013; Struth, 2019).

At any rate, the German population currently seems rather skeptical about the state of freedom of speech in their country: According to a survey conducted by the Allensbach Institute for Public Opinion Research in 2021, the number of citizens who believe that it is better to be cautious when voicing political opinions in Germany has reached an all-time high with 45% of the population agreeing (Petersen, 2021). Although research is only starting to examine the reasons for these perceptions, the results seem to indicate that, first, almost half of the citizens have experienced or at least heard of instances in which voicing certain political opinions had negative consequences. Second, these findings might also indicate an insecurity or lack of consensus about what can be said freely and the kind of statements that might trigger a backlash of critical remarks, counter-speech, or even hate—justifiably or not. However, this does not seem to be the case for all kinds of statements. Recent findings show that while the evaluation of some statements that have been the subject of public debate diverges significantly among the population, there is quite a large consensus on others (Petersen, 2021). Thus, it seems that perceptions vary based on the statement.

Such differences can not only be seen in the general population but also in the judiciary, where they are even more consequential because they

result in divergent legal assessments of potentially justiciable statements by different judicial bodies. Prominent examples in Germany include differing verdicts on insulting statements against well-known German politicians (e.g., Berlin District Court, 2020; Hong, 2020) and diverging evaluations of political campaign posters of a right-wing extremist party (e.g., Kister, 2021). However, while these examples show the complexity of the topic even in court, the discussions about the judgement of statements usually begin well before a possible legal dispute and often do not occupy the courts. In most cases, the question is not whether statements violate legal norms but whether they violate social norms by, for example, making discriminatory claims. Therefore, given recent heated debates, it is important to investigate where the boundaries lie between the "sayable" and "unspeakable" for citizens, how much of a consensus there is about these classifications, and what individual-level factors might explain differing assessments of statements. The answers to these questions are addressed in the present article.

Moreover, the issues we investigate here are socially relevant, especially because the lack of a minimum consensus on the social acceptability of speech can be seen as a danger to social integration (e.g., Quiring et al., 2020). If one segment of society is under the constant impression that its freedom of speech is restricted while another feels that it is constantly belittled or insulted, then feelings of deprivation and social distrust, affective polarization, or even social intergroup conflicts are potential consequences.

In this study, we start with two assumptions. First, we assume that the perception of what can be considered a socially acceptable statement varies between individuals (e.g., Mummendey et al., 2009) and that individual predispositions play a central role in explaining those differences (e.g., Valkenburg & Peter, 2013). For example, in the context of motivated reasoning, numerous studies on information processing show that attitudes, identities, and values have a considerable influence on how information is perceived and interpreted (e.g., Kahan, 2013; Taber & Lodge, 2006). Further, findings from research into polarization suggest that statements about a certain group are perceived differently by group members and non-members and that outgroup derogation can increase attitude polarization (Wojcieszak et al., 2021).

Second, we assume that individual differences are not the same across statements. As shown in previous research, there are statements, terms, and expressions that are likely to be either rejected or accepted relatively uniformly, while there probably is less consensus in terms of others (Petersen, 2021). This could apply, for example, to more subtle forms of

discrimination, which have also been discussed for some time as "microaggressions" (e.g., Lilienfeld 2017; Sue, 2010; Torino et al., 2018). Dissent over the evaluation of such statements could indicate that these evaluations are changing because, for example, social power relations are shifting. Furthermore, it could also be due to interest groups (from less powerful or marginalized groups) articulating and problematizing their critical views on statements and forms of speech that were previously considered "unproblematic" by a majority or the actors dominating public discourse. It is these kinds of developments that point to the necessity for democratic societies to find a minimal common ground at least about what can be regarded as discriminatory speech in order to strengthen respectful and inclusive public discourse.

Against this background, we want to explore the actual degree of social consensus in the assessment of potentially controversial and harmful statements and what individual characteristics might help explain possible individual-level differences. To this end, we conducted a quantitative survey, in which we presented respondents with a set of statements and asked whether they considered these statements as acceptable or unacceptable and as hurtful or not. Given that gender plays a central role in victimization through hate speech and in the debates about non-discriminatory speech, we decided to use statements about women as examples, among which there were those that could be regarded as discriminatory, hateful, and therefore (potentially) harmful.

Freedom of Speech and (Potentially) Harmful Speech

One of the starting points for this paper is the idea that public controversies about the social acceptability of certain statements might be a reason for the widespread impression that freedom of speech is increasingly restricted. This does not mean, of course, that a pluralistic democratic society should aim for a situation in which no such controversies exist. This is neither desirable nor realistic. Some scholars also argue that perceptions of "taboos" and perceived social restrictions to speech are entirely normal and even necessary because even a free and democratic society has to rely upon at least some consensus on speech norms (e.g., Quiring et al., 2020). Nevertheless, we argue that it is important to understand how and why citizens differ in their perceptions of speech norms, as these might turn into more fundamental doubts about the ability to voice one's opinions and maybe even about the functioning of democracy per se. At the same time, if freedom of speech is indeed restricted by state authorities, Internet

companies, or a toxic culture of hateful discourse, this might also pose a threat to democracy (e.g., Quiring et al., 2020).

Although not much research has been published on this issue, referring to the case of Germany, public interest in perceptions of freedom of speech and of specific statements has been on the rise in recent years. In fact, public debate about "political correctness" and "cancel culture" in the country has also been driven by polls published by the Institut für Demoskopie at Allensbach (IfD). Although the institute has long been interested in "taboo issues", "political correctness," what is "sayable," as well as related subjects for decades—in the footsteps of its founder Elisabeth Noelle-Neumann—public interest has been especially intense since seemingly alarming results were published in 2019 and again in 2021 (Köcher, 2019; Petersen, 2021). For example, the most recent study found that just as many people felt that one should "rather be careful" in expressing political opinions (44%) as those that held the view that one can "express one's opinion freely" (45%) (Petersen, 2021, p. 22).

Besides data for general perceptions of freedom of speech in Germany, the IfD has also been exploring contested and sensitive issues (Noelle-Neumann, 1996; Petersen, 2013; Köcher, 2019). Results show that topics such as "foreigners," "asylum seekers" and "refugees," "Muslims/Islam," "Judaism/Israel," or Germany's National Socialist past are among the issues that have been perceived as sensitive for decades (Köcher, 2019, p. 15). In addition, the IfD has been asking respondents for their perceptions of more concrete political statements (e.g., "Refugees are criminal"), examples of non-discriminatory speech that were sometimes labelled as being a part of a trend toward "political correctness" (e.g., Köcher, 2019, pp. 15–17; Petersen, 2021), and statements that could be viewed as "especially sensitive." Respondents were asked, for example, to decide whether such statements should be forbidden or whether they could get into hot water for the same (e.g., Petersen, 2013). However, although these data are interesting, the published findings are confined to descriptive aggregate-level analysis for the most part. Nevertheless, more in-depth systematic analyses of the reasons for these perceptions are missing. This includes the characteristics of the statements and issues that should be judged or the individual-level characteristics of citizens that drive these perceptions. So, the questions of who perceives certain potentially controversial statements in a particular way and why this is the case have remained largely unanswered.

Till date, the only German study to date that has provided a more in-depth analysis of the aforementioned kind is a recent investigation into perceived "speech bans" (Quiring et al., 2020). Taking four contested issues as examples, they investigated whether restrictions on free speech

were perceived in relation to them, how this perception was connected to personal opinions on those issues, and who was most likely to perceive speech restrictions on those issues (Quiring et al., 2020, p. 61). In line with assumptions derived from Noelle-Neumann's work, they found that restrictions on freedom are more likely to be perceived in terms of morally loaded issues and those on which the public opinion is split (in this case, religion, migration, and criminality). In addition, they found the dissonance between an individual's opinion and that of the (perceived) majority makes the perception of restrictions more likely (Quiring et al., 2020, pp. 67–68). Furthermore, women and respondents who are less satisfied with democracy, have a lack of trust in traditional media, and are less socially integrated tended to perceive more restrictions on speech about these contested issues.

While Quiring et al. (2020) investigated the perceptions of speech bans for specific political issues, we set a slightly different focus here by investigating the perception of everyday utterances that people might encounter more generally among personal contacts, on social media, and in online comments sections. In addition, we do not ask about "speech bans" but whether people perceive certain statements as acceptable and potentially hurtful. This might change the perspective of respondents to a certain degree by putting a greater focus on those who may be negatively affected by a speech act and by not implying that there may be actors, institutions, or powers that would be able to "order" a speech ban. However, the fundamental question remains the same: Is there consensus or disagreement in the perception of statements and what drives potential differences?

Freedom of Speech in the Context of Misogynistic Statements

While free speech is discussed in relation to various topics, we examine the issue using the example of misogynistic speech. This topic was chosen for two reasons. The first is because of the continued relevance of the research topic: despite increasing awareness of gender equality, the sexist treatment of women in the form of stereotypically derogatory expectations and expressions is still present as shown by recent studies (e.g., Lui & Quezada, 2019; Foster, 2009). In the context of microaggressions, Sue (2010) even referred to gender as probably the most restrictive force in everyday life (p. 160ff). In addition, women run a higher risk of being victims of online hate speech (e.g., Chen et al., 2020).

The second reason is that it has been shown that there might be a gender difference in the perception of freedom of speech. A study at

American universities, for example, asked students about their assessment of the relevance of an unrestricted right to freedom of speech. When seeing gender as a binary concept, the results showed apparent differences between men and women. Nearly 60% of women thought an inclusive, open society was more important than freedom of speech as compared to only 28% of men (Knight Foundation, 2019, p. 6). Nearly half of female students also said Americans need to be more careful in their own choice of words, while only 26% of men agreed. In contrast, 74% of men were of the opposite opinion, saying people too often overreact to statements (p. 8f). In addition, 53% of women thought that hate speech should not be protected by the First Amendment, while 74% of men disagreed (p. 10). Even though these findings are restricted to the American context, gender may also impact the aforementioned perceptions in other countries such as Germany.

Types of (Potentially) Harmful Speech

Before investigating individual characteristics that may lead to different perceptions of controversial, (potentially) harmful statements, it is crucial to first define the characteristics of these statements. Previous research usually did not distinguish between different types of speech (e.g., Petersen, 2021; Wegner et al., 2020) or only focused on the perception of one specific type of discriminatory language, such as hate speech or microaggressive speech (e.g., Sue, 2010), without connecting it to the larger picture of freedom of speech. Even if the discussions surrounding certain controversial statements do not exclusively revolve around the juxtaposition of freedom of expression and freedom from discrimination, the discourse can mostly be traced to the controversy between these two fundamental freedoms (e.g., Struth, 2019). In order to get closer to understanding the perception of this conflict in the general population, in this study, we try to distinguish different degrees of discriminatory statements against women based on their extremity and the blatancy of the insult they represent. In doing so, we refer to the definition of discrimination by Dovidio et al. (2010), who defined it as behavior that "creates, maintains, or reinforces advantage for some groups and their members over other groups and their members" (p. 10).

One type of potentially discriminatory language that has been increasingly discussed in recent years, especially in the context of political correctness, is that of *microaggressions*. It refers to an implicit and subtle devaluation of discriminated social groups that can be verbally or behavio-

rally elicited (Sue et al., 2007). The concept of microaggressions describes insults that are less obviously recognizable as disparagement than other forms of discrimination. However, at the same time, they are in no way considered to be less derogatorily motivated (Lilienfeld, 2017, p. 139). With this type of potentially harmful language in particular, subjectivity plays a significant role. In order to label microaggressions as such, the individual assessment of the person affected is required: "First, the person must determine whether a microaggression has occurred" (Sue et al., 2007, p. 279). Moreover, the affected person does not necessarily have to assume that the communicator intended to offend. Instead, the pejorative may have also been uttered unconsciously (Sue et al., 2007, p. 278). This highlights that microaggressions, in particular, are in the eye of the beholder according to current research, and it depends on the individual whether a statement can be considered problematic at all. This means that not even all members of the respective social group may perceive microaggressive speech as discriminatory or hurtful (e.g., Sue, 2010; Lilienfeld, 2017)

In contrast to subtle microaggressive speech, is blatant discrimination in the form of hate speech. It can be understood as speech "that involves the advocacy of hatred and discrimination against groups on basis of their race, colour, ethnicity, religious beliefs, sexual orientation, or other status" (Boyle, 2001, p. 489). In the present research context, the victims of sexist hate comments are women. Moreover, the difference between hate and microaggressive speech becomes clear in terms of the perceived communicator's intentionality. Here, the intention to offend others is not questioned by the person affected by the hate speech (Marker, 2013), especially since the explicit use of offensive language also shapes it. In terms of perception, studies have found that hate speech is recognized as such by the majority of people and that it is perceived as disturbing, although this also depends on various subjective factors such as gender (e.g., Costello et al., 2019).

As microaggressive and hate speech represent extreme points of (potentially) hurtful speech in terms of the blatancy of discrimination, we will also include an intermediate type of speech in the present study. This will help us examine grey areas within the spectrum of speech as well. We will call this form of discriminatory, that are potentially harmful and controversial statements, "derogatory speech." It describes a more overt expression of discrimination than microaggressive speech and differs from hate speech in so far as the choice of words is less offensive and violent.

Thus, in the following, we distinguish three types of speech: microaggressive, derogatory and hate speech. With regard to the perception of this spectrum, research has so far been devoted, for example, to the psychologi-

cal consequences after reception (for example, mental illnesses as a result of hate messages; for e.g., Leets, 2002). Moreover, comparing the different forms, some literature suggests that these psychological consequences are even worse for more subtle statements, such as verbal microaggressions (e.g., Williams & Mohammed, 2013; Sue, 2010). However, we know relatively little about a concrete comparison of perceptions of different forms of discriminatory speech (Lui & Quezada, 2019), especially in the context of freedom of speech debates. This research gap will therefore be addressed in the present study.

Predictors of the Perception of (Potentially) Harmful Speech

As argued above, we assume that individual characteristics impact the way potentially controversial, microaggressive, derogatory, or hateful speech against women is perceived. In this analysis, we will test four sets of potential predictors.

The first set of predictors are sociodemographic factors that can also be regarded as being indicative of different social identities. This is, of course, especially true for gender. It almost seems self-evident to assume that women will perceive derogatory and other problematic or controversial statements about women as less acceptable than men. In fact, research shows that gender does affect perceptions and attitudes on a gender-related issue like gender-neutral speech or gendered job announcements (e.g., Budziszewska et al., 2014; Gustafsson Sendén et al., 2015). Moreover, as noted earlier, it was found that women more often prioritize inclusive, cautious language over free speech (Knight Foundation, 2019) and more frequently perceive hateful language as disturbing than men (e.g., Costello et al., 2019). We, therefore, put forward our first hypothesis below.

H1: Women will perceive microaggressive, derogatory, and hateful statements against women as less acceptable than men.

In addition, it can be assumed that the way women are addressed and talked about has been the subject of social change over the last decades that have seen processes of emancipation and growing societal awareness of questions of gender equality (e.g., Inglehart & Norris, 2003; Scarborough et al., 2019). Consequently, we can assume that younger generations have grown up and socialized in an environment that has become much more positive with respect to gender equality. Therefore, we assume that younger people should also be more critical of discriminatory speech against women. We, therefore, put forward our second hypothesis below.

H2: Members of younger generations will perceive microaggressive, derogatory, and hateful statements against women as less acceptable than members of older generations.

The final sociodemographic factor we investigate is education. Generally, it can be shown that more educated people hold less traditional and more progressive values. Further, they have also been shown to be less sexist, although it has to be stressed that sexist attitudes are, of course, not restricted to the less formally educated (e.g., Pew Research Center, 2019). In addition, formal education has been shown to correlate with a stronger awareness and preference for gender equality (Pew Research Center, 2019). As we have not included direct measures of traditional and progressive values, sexism, and attitudes towards gender equality, we therefore view formal education as a proxy for these in the context of this study. In addition, it can also be argued that less formal education may also have a more direct effect on perceptions of speech because it may correlate with a more frequent usage of harsh language, although this is a topic of debate in linguistics (e.g., Love, 2021). We therefore assume the following.

H3: More formally educated people will perceive microaggressive, derogatory, and hateful statements against women as less acceptable than people with less formal education.

The second set of factors we consider here are experiences of discrimination based on, for example, migration background, sexual orientation, gender, etc. Discrimination, in general, may not only result in anger, anxiety, stress, or even mental health problems, but people experiencing verbal or non-verbal discrimination may also perceive future social interactions differently due to their experiences (e.g., Mummendey, 2009). One potential effect may be that people become increasingly aware of and sensitive to discrimination in general and discriminatory speech in particular and thus regard even more subtle forms of microaggressive speech as less acceptable. In contrast, extreme forms of non-verbal discrimination in particular may also raise the bar for verbal discrimination to be perceived as "real" discrimination ("I've experienced worse.") and thus lead to a de-sensitizing effect. However, as the literature mostly suggests that it is more likely that experiences of discrimination tend to raise awareness and make people more sensitive towards it, we assume the following.

H4: The more that people have experienced discrimination, the less acceptable they will find microaggressive, derogatory, and hateful statements against women.

In addition, this effect of discrimination should be especially pronounced in the case of statements against women for people who have been discriminated against based on their gender. We therefore assume:

H5: The more that people have experienced discrimination based on their gender, the less acceptable they will find microaggressive, derogatory, and hateful statements against women.

The third set of factors we include are political predispositions. Generally, we can assume that politically more conservative and right-wing individuals will have more traditional values and preconceptions about gender roles and be more critical of gender equality in general and non-discriminatory speech in particular (e. g., Christley, 2021). In addition, they may be more inclined to be critical about issues around "political correctness" and what they perceive as dangers to freedom of speech, as recent research indicates (e.g., Petersen, 2021). Moreover, concerning hate speech, some studies have already shown that the perception of certain statements varies depending on how left or right people locate themselves on the political spectrum (e.g., Costello et al., 2019). This may be especially true for voters of the right-wing populist Alternative for Germany (AfD), which has—like other right-wing populist parties—positioned itself as a strong critic of progressive gender-equality policies (e.g., Abi-Hassan, 2017; Petersen, 2021). In contrast, people who are politically more progressive and those on the left of the political spectrum can be regarded as more sensitive toward gender-equality issues. We therefore assume the following.

H6: The more left-wing political attitudes people have the less they will perceive microaggressive, derogatory, and hateful statements against women as acceptable.

H7: Voters of right-wing populist AfD will perceive microaggressive, derogatory, and hateful statements against women as more acceptable than voters of other parties.

Finally, we also include trust in traditional media and trust in online user-generated content as potential drivers of differences in the perception of discriminatory speech. We opted for these indicators because trust in media can be a better predictor of certain media effects than media usage itself (e.g., Fawzi et al., 2021). This is because the processing and interpretation of information are strongly affected by the trust people have in a source and because not all people using a certain source actually regard it as trustworthy (Strömbäck et al., 2020). In addition, we argue that trust in different types of media sources may be connected to differing perceptions of controversial statements about women because it also reflects a certain degree of agreement with the basic values apparent in the content of these media. Against this backdrop, we first assume that most German traditional journalistic sources tend to position themselves against discrimination and more in favor of gender equality—although there are definitely exceptions and differences. Second, we assume that

user-generated content online tends to contain more discriminatory, hateful, and misogynistic speech than the content of traditional journalistic news sources. Therefore, people trusting user-generated content might also be more likely to regard such speech as more acceptable because they are more often confronted with this kind of speech, might get used to it, and thus regard it as more appropriate ("normalization"). We, therefore, put forward our last hypothesis as follows.

H8: The higher their trust in traditional media, the less will people perceive microaggressive, derogatory, and hateful statements against women as acceptable.

H9: The higher their trust in online user comments and posts, the more will people perceive microaggressive, derogatory, and hateful statements against women as acceptable.

Methods

To test our hypotheses, we conducted a quantitative online survey. It was part of a research seminar at the Department of Media and Communication at LMU Munich. With the help of an access panel from a market researcher (*Dynata*), a quota sample was drawn within the age range of 18 to 65 years of the German voting population. Quotas were applied for gender, age brackets, and education. After intense pretesting, the survey was conducted in January 2020. Straightliners and respondents with an interview duration of less than six minutes were excluded from the sample. In addition, due to the very small group size, respondents who classified themselves as non-binary when stating their gender were excluded from the analysis (N = 2). Overall, 20% of original respondents were excluded for quality assurance purposes resulting in N = 943.

However, the questionnaire contained a split-ballot section on the perception of discriminatory statements. One half of the sample was presented with statements against people with a migration background and the other with statements against women. Since the latter constitutes the core of the present paper, the following section will assume that there were only half of the subjects. This results in a sample size of N = 447 respondents for the descriptive analyses and N = 401 for the explanatory analyses.

Dependent Variables

The core of this study is the evaluation of statements against women. To this end, the respondents were presented with a series of comments against women, which corresponded to the spectrum of forms of expression previously described: two statements each that can be assigned to the levels of microaggressive, derogatory statements, and hate speech; two neutral statements that we regarded as non-discriminatory for comparison purposes.

The conceptualization of these statements was based on previous research on discriminatory language against women (Sue, 2010). Here, such statements were defined based on their sexualization and objectification of women or viewing them as less competent and intelligent. Based on these preliminary considerations, actual comments from Internet forums were used for the operationalization and subsequently adapted to the respective levels. The statements not only differed in their extremity but also in the choice of words, which explicitly distinguishes the two levels of derogatory and hate speech.

Respondents also saw two neutrally phrased control items. This resulted in a total of eight items, which were randomized and presented to the respondents twice for evaluation. First, respondents were asked to evaluate the respective statements more rationally by asking whether they considered the items acceptable or thought they went too far. This was measured in each case on a seven-point scale (1 = "is acceptable", 7 = "goes too far"; the "don't know" option was also available). The second question was aimed at assessing the potential of the statement to elicit a negative emotional response by a potential receiver. Here, respondents were asked whether they thought that the statements could be hurtful ("Would you classify these statements as hurtful or would you say they are not hurtful?"; 1 = "not hurtful," 7 = "hurtful"; the "don't know" option was also available). Both questions were immediately asked one after the other (Table 1)[1].

1 The conceptualization and operationalization of the statements were undertaken by Danilo Harles, Lilli Fischer, Velina Chekelova, and Anna-Luisa Sacher.

Table 1: Set of (Potentially) Harmful Statements Against Women

Perception of acceptability		Perception of potential to hurt
"Now it is about very specific statements: Imagine the following comments are made to a woman. Do you find these statements acceptable, or do you think they go too far?"		"Below, you find the comments from the previous question again. Regardless of whether you find them acceptable in principle, would you classify these statements as hurtful, or would you say they are not hurtful?"
Neutral speech	"Are you free for a short meeting tomorrow?"	
	"Have you seen, it is really nice weather today."	
Micro-aggression	"You must be on your period, right?"	
	"I think that is brave, though, that you have a career on top of having kids."	
Derogatory Speech	"Well, of course, you earn less, this should be the case for you women."	
	"You women just do not belong in the office, you should be taking care of the household."	
Hate Speech	"Stupid and ugly, it takes a woman's quota for you to get a job."	
	"If you dress like that, don't be surprised if you get raped."	

Independent Variables

Sociodemographics

Next to gender (47.5% female; 52.1% male), another sociodemographic factor of interest is age (M = 48.21, SD = 12.26), which is also measured using generations ("Baby Boomers": 1945–1964; "Generation X": 1965–1981; "Generation Y": 1982–1994; "Generation Z": 1995–2010). Since our sample is limited to the age range of 18 to 65, the generations under study are also limited to 1954 to 2001. In addition, respondents' education was obtained, and the variable was dichotomized to indicate whether they had received a high school diploma or not (1 = no diploma, 42.8%; 2 = diploma, 57.2%).

Experiences of Discrimination

Concerning previous experiences of discrimination, respondents were presented with nine items on a five-point scale to indicate how often they had already been disadvantaged or discriminated against for a variety of reasons (1 = never, 5 = very often). Potential reasons mentioned were, for example, migration background, appearance, sexual orientation, and gender. For the following analyses, this variable is relevant in two ways. First, the influence

of "experiences of discrimination based on one's gender" is considered individually (M = 1.63; SD = 1.037). Second, a mean index was formed from all items (M = 1.48; SD = 0.66; α = .876) in order to make a statement about the effect of "the variety of previous experiences of discrimination." Respondents with more frequent and diverse experiences will score higher on this index.

Political Predispositions

In order to also examine the influence of political attitude on the perception of statements, two factors were considered. First, we examined "general political positioning" using an 11-point left-right scale (1 = left, 11 = right; M = 5.87, SD = 2.04). Second, we measured "party preference" using the so-called Sunday question, in which respondents indicated which party they would vote for if an election were held next Sunday (CDU/CSU: 16.1%; Green Party: 18.7%; AfD: 11.5%; SPD: 9.2%; FDP: 6.7%; The Left: 9.8%; others: 5.2%).

Media Trust

Finally, subjects' "media trust" was measured by four items in one question. Due to very high correlations between two indicators ("User generated commentary and posts online" and "Social media") and more than 200 "don't know" responses for another item ("Alternative media"), we only included two of the items in the analysis. Here, respondents were asked to rate their trust using a five-point scale (1 = no trust, 5 = very high trust) regarding traditional media (e.g., newspapers, news magazines, radio, television; M = 3.65, SD = 1.14) and posts or comments by Internet users (e.g., in forums, blogs, or comment sections; M = 2.17, SD = 1.05).

Results

Descriptive Analyses of Perceptions of (Potentially) Harmful Statements

The first result of our analysis is, rather surprisingly, that respondents perceived the statements very similarly no matter whether they were asked to evaluate their acceptability or hurtfulness. We only found a slight

deviation in the perceptions of the statements. The distinction between the more rational assessment of whether the statements are acceptable and the more emotional assessment of whether they are hurtful does not seem relevant to respondents. This may be caused by the fact that the two assessment dimensions are interwoven and thus cover the same construct: if one classifies a statement as acceptable, this may also mean that one would interpret the degree of violation as low and vice versa. However, this could also be due to a response bias in the form of consistency effects since respondents were presented with the two item batteries immediately, one after the other. Thus, they possibly rated the statements consistently on the five-point scale despite different types of randomization. Since the evaluations of the more rational and emotional assessment dimensions were so similar, the results will be considered together as an index for each level in the following analyses. Reliability coefficients and index means are documented in Table 2.

Second, neutral control statements were indeed overwhelmingly perceived as acceptable and not hurtful (M = 1.88; SD = 1.26). However, it is interesting to note that the statement "Are you free for a short meeting tomorrow?" was rated as somewhat more hurtful and unacceptable (M = 2.14; SD = 1.51) than the statement about the weather (M = 1.62; SD = 1.29)—and as expected from a neutral item. Here, respondents possibly interpreted this statement in the context of the discriminatory comments as a courtship towards women and thus did not understand it as completely neutral. The difference is nevertheless small.

In terms of microaggressive speech, there was a striking difference between the two items as shown below. The statement "I think it's brave that you have a career while raising children" was rated below the midpoint of the scale for both dimensions (M = 3.34; SD = 1.95). However, respondents considered the item "You certainly have your period, don't you?" to be less acceptable, with the mean being way above the midpoint (M = 5.69; SD = 1.58). A comment regarding a woman's role image that it is courageous to be employed while raising children seems to be socially viewed as more "sayable" than a statement referring to her period, which is likely perceived as a violation of a woman's sphere of intimacy.

Concerning the results of the two extreme discriminatory levels, derogatory language and hate speech, it was found that all four statements were rated similarly. Accordingly, it did not seem to make a difference whether the statements were provided with extreme wording, as with hate speech (e.g., "Stupid and ugly … it takes a woman's quota for you to get a job"). The apparent disparagement provided a perceived boundary-crossing for all items at both levels. Overall, the theoretical differentiation proved to

not be completely in line with the respondents' perception. This is especially true for the derogatory and hate speech statements, with the former being perceived as even more unacceptable than one of the statements we had classified as hate speech (Table 2).

Table 2: Perceptions of (Potentially) Hurtful Statements (M, SD, Indices)

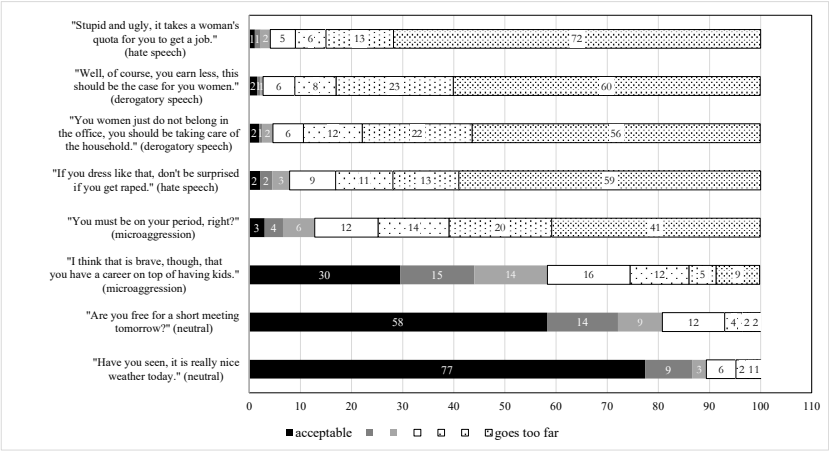
Neutral speech	"Are you free for a short meeting tomorrow?"	M = 2.14 SD = 1.51	M = 1.88 SD = 1.26 $r = .564^{***}$
	"Have you noticed the really nice weather today?"	M = 1.62 SD = 1.29	
Micro-aggression	"You must be on your period, right?"	M = 5.69 SD = 1.58	M = 4.51 SD = 1.44 $r = .264^{***}$
	"I think that it is brave, though, that you have a career on top of having kids."	M = 3.34 SD = 1.95	
Derogatory speech	"Well, of course, you earn less … this should be the case for you women."	M = 6.38 SD = 1.11	M = 6.32 SD = 1.09 $r = .685^{***}$
	"You women just do not belong in the office … you should be taking care of the household."	M = 6.26 SD = 1.22	
Hate speech	"Stupid and ugly … it takes a woman's quota for you to get a job."	M = 6.49 SD = 1.12	M = 6.29 SD = 1.11 $r = .425^{***}$
	"If you dress like that, don't be surprised if you get raped."	M = 6.09 SD = 1.43	

Note. Based on N = 401 respondents. Spearman-Brown r, *** p < .001

One of the questions we started with was whether there is a consensus about assessing potentially harmful statements about women among German citizens. Figure 1, which contains descriptive results for the acceptability assessment, suggests that this is mostly the case at least for the statements we classified as hate and derogatory speech. Moreover, only between 2% and 7% of respondents consider these statements to be more or less acceptable, which itself is rather surprising. Meanwhile, the consensus is almost as high for the statement regarding a women's period, although the share of respondents choosing the extreme point of complete acceptability is, in fact, lower. Then, the most diverse responses were regarding the statement that addresses the conflict between having kids and a career. While just more than half of the respondents find the statement rather acceptable, this is not the case for almost another 30%. Obviously, this

statement could be a trigger for controversy, probably be hurtful for a least some women, or touches a hot topic that is still a social taboo (Figure 1).

Figure 1: Distribution of Answers – Perceptions of (Potentially) Harmful Statements



Note. Based on N = 401 respondents.

While the grouping of the statements examined was based on theoretical assumptions, the results indicate other clustering and thus the items were subsequently examined using an exploratory factor analysis (Varimax). First, the Kaiser-Meyer-Olkin criterion was fulfilled (KMO=.76), and Bartlett's test was found significant (χ^2 (28) = 1174.96, p< .001). Here, two factors can be extracted, each of which has an Eigenvalue greater than 1. They support the assumptions described above: One factor covers the neutral statements and the microaggression regarding a woman's career choice. The second factor, in contrast, includes the second microaggression regarding the female menstrual cycle, which was considered to be less acceptable, as well as the statements representing the more extreme types of speech. Accordingly, the statements we regarded as microaggressions were not perceived similarly by our respondents. While one resembles a neutral item more closely, the second was perceived almost in the same way as derogatory and hate speech. Together, the two factors can explain 60.58% of the variance.

Explanatory Analyses of Predictors of the Perceptions of (Potentially) Harmful Statements

We ran four regression models to explain individual perceptions of the different types of statements. However, the variance explained by the models was rather small and did not differ much, with a range between 10% and 13%. Of course, this is not that surprising given that the overall variance of perceptions is somewhat limited, especially in the case of neutral, derogatory, and hateful statements. In addition, the number of cases in the analysis was reduced to N = 403 due to missing values in the questions on media trust (Table 3).

The first remarkable finding is that only in one case did gender make a difference in perceptions when controlling for the other included factors. Further, only in the case of hate speech did women perceive the statements as slightly less acceptable and harmful than men. H1 is therefore rejected for most types of statements. As for age, a more consistent picture across types of statements is apparent, but it is rather surprising: most notably, members of generation X perceive all types of statements as more acceptable than the reference group of Boomers. For the neutral statements, this is also true for generation Y. In contrast, there mostly are no differences between the younger generations Y and Z and the Boomers. Therefore, H2 has to be rejected as only generation X seems perceive the statements as more acceptable. As for education, we only see a small effect in the case of the neutral statements, with higher education even contributing to a less critical view of these statements. H3 is therefore rejected. More generally, the explanatory power of sociodemographics is rather limited in our analysis, with age showing the most consistent impact.

The picture is slightly different for our indicators of experiences of discrimination that show a rather complex pattern of influences. For neutral statements, we found a significant effect of our cumulative measure of various types of discrimination with various experiences contributing to a more critical perception of the neutral statements. Meanwhile, microaggressive speech had different results. Respondents who have experienced various types of discrimination perceived these kinds of statements as "more" acceptable, whereas those who have experienced discrimination with respect to their gender regard these as "*less*" acceptable. However, this result has to be interpreted cautiously because, as we have seen, the two statements combined in the index were evaluated rather differently. For the derogatory statements, none of the indicators produced significant effects, but the cumulative indicator pointed in the same direction as that for microaggressions. Finally, the same indicator showed a significant

effect on the perception of hate speech with more diverse experiences of discrimination again contributing to a less critical stance towards hate speech. Overall, these results suggest a sensitizing effect of gender-specific discrimination concerning microaggressions and a de-sensitization or habituation effect for more varied experiences of discrimination in terms of microaggressive and hate speech, with the coefficient for derogatory speech at least pointing in the same direction. Accordingly, H4 has to be completely rejected, and H5 has to be mostly rejected because it can only be confirmed for microaggressive speech.

Moreover, only a small number of significant effects can be identified concerning political attitude factors. The more conservative the respondents ranked themselves on the left-right scale, the less negative they perceived the microaggression, while there was no effect for the other types of speech. That means that we can only partially confirm H6. With regard to party preferences, we can also confirm our assumptions only partially: respondents with voting intentions for the AfD considered derogatory statements to be more acceptable. This means that we can also only partially confirm H7. Meanwhile, other effects were only found in one other case, with a preference for the FDP resulting in neutral statements being rated as less acceptable.

Finally, there were several significant effects in terms of the two trust indicators. As it turns out, neutral, microaggressive, and derogatory statements were perceived as less acceptable by respondents who tend to have a higher trust in traditional media. Therefore, this is in line with H8, which can be mostly confirmed. The picture is less clear for trust in posts and comments from Internet users. While neutral statements were regarded as less acceptable for those who have a higher trust in user posts and comments, derogatory speech was perceived as more acceptable by those with a higher trust in user posts. In addition, no effects could be found for microaggressive and hate speech. Therefore, H9 has to mostly be rejected (Table 3).

Table 3: Predictors of the Perception of (Potentially) Harmful Speech

Predictors	Neutral statements *Beta*	Micro-aggression *Beta*	Derogatory Speech *Beta*	Hate Speech *Beta*
		Sociodemographics		
Gender (0 = male. 1 = female)	-,030	-,024	,067	,101^{+}
Gen. Z (18-25 years)	-,064	-,056	-,020	-,016
Gen. Y (26-39 years)	-,130*	,011	,072	,008
Gen. X (40-54 years)	-,122*	-,105*	-,107^{+}	-,109*
High School Diploma (0 = no. 1 = yes)	-,112*	-,004	-,081	,028
R^2	*.029*	*.022*	*.027*	*.032*
		Experiences of discrimination		
Discrimination based on gender	-,007	,235**	-,065	,035
Discrimination in general	,179*	-,177*	-,106	-,244***
R^2	*.063*	*.049*	*.065*	*.089*
		Political predispositions		
Political attitude (left-right scale)	-.006	-,156*	-.019	.006
Party preference (0 = no. 1 = yes)				
CDU/CSU	,082	,062	-,018	,006
Green party	,049	,002	-,042	,104
AfD	,036	-,012	-,150*	-,104
SPD	,049	,012	-,060	-,036
FDP	,122*	,019	-,030	-,052
The Left	,051	-,079	,043	,082
R^2	*.085*	*.087*	*.106*	*.128*
		Media trust		
Traditional media	,110*	,138*	,159**	,046
Posts/comments by Internet users	,165**	,061	-,102*	,011
R^2	*.118*	*.104*	*.135*	*.130*

Note. Table entries are beta coefficients from linear regression analyses. Based on N = 401. * p < .05; ** p < .01; *** p < .001

Conclusion

The starting point of this paper was the notion that it is important to understand where people draw the line between acceptable and unacceptable speech and what drives these perceptions. This is because differences in the perception of (potentially) discriminatory speech are at the heart of the frequently recurring disputes about controversial statements. Further, these controversies are important because, at least in Germany, they are potentially one factor contributing to the widespread perception that freedom of speech has become increasingly restricted. Against this background, we asked two rather basic questions: (a) whether and to what extent people differ in their perception of (potentially) harmful speech against women; and (b) which individual-level characteristics might explain such differen-

ces. The results of our analyses that are based on a survey conducted in early 2020 can be summarized as follows:

(1) Our respondents' answers did not vary much when they were asked to evaluate the acceptability and harmfulness of statements. Whether this is a methodological artefact because of consistency effects or whether this indicates that the acceptability of statements is to a large extent driven by the evaluation of their harmfulness cannot, however, be decided on the basis of our data.

(2) Most of the statements that we came up with based on real Internet comments and that we asked our respondents to evaluate were perceived rather similarly. This means that there was a large consensus about their evaluation. This is especially true for the statements against women that we had previously classified as derogatory and hate speech. Here, no more than about 10% of respondents regarded them as more or less acceptable. Our theoretical distinction between the two groups based on particularly offensive language was not supported by the results; the factor analysis subsumed them into one factor. Accordingly, our results confirm findings from previous studies that offensively discriminatory statements are also recognized as such by a majority and are thus strongly rejected (e.g., Costello et al., 2019). Although this might sound like good news, it should be noted that even a small number of people can make a difference in (online) discourses. In this context, given the extremity of the statements, there is still worry about the number of people viewing them as acceptable.

(3) The statements we had classified in advance as microaggressions were evaluated very differently. This was already evident based on the descriptive statistics but became apparent in the subsequent exploratory factor analysis, which did not support our grouping based on theoretical assumptions. Instead, two factors were distinguished that separated the two microaggressions based on the results. One statement was almost rated as unacceptable as the derogatory and hate speech statements, while the other was perceived in more diverse ways, but closer to the rating of the neutral statements. The obvious discrepancy between the a-priori classification and the respondents' perception is interesting in itself and highlights the difficulty of determining in advance how some statements may go down with an audience. In addition, the fact that there was no consensus on one of the statements reflects the more subtle and unclear nature of microaggressions that might result in controversy and misunderstanding because of entirely different perceptions and interpretations. Therefore, the central role of the subjective views on microaggression, which has been discussed in previous studies (e.g., Sue et al., 2007; Sue, 2010), is also

reflected in our results. Moreover, they also point to potential conceptual challenges with the construct of microaggression that remain to be resolved in future research (e.g., Lilienfeld, 2017).

(4) The individual characteristics we included did not explain much of the (mostly) rather small perceptual differences we encountered. This is particularly interesting in terms of the individual affectedness based on respondents' own gender being discussed. In contrast to our expectation that women's direct involvement might result in different perceptions, there were few to no differences. Thus, in terms of perceptions of discrimination against women, group membership did not play a central role in our study, so we could not replicate previous findings (e.g., Knight Foundation, 2019; Costello et al., 2019). However, this could also be due to methodological reasons, which we will discuss later. In this context, since there are certainly structural differences between discriminated groups, replications with other target groups would be of interest.

(5) Considering the other predictors analyzed, the most consistent effects across different types of statements appeared to come from age and trust in traditional media. In general, members of Generation X found the statements somewhat more acceptable than all other generations, younger and older. The reasons for this unexpected finding are unclear and will have to be identified in further analyses. The consideration that younger people have been socialized more sensitively due to increasing awareness of gender equality (e.g., Inglehart & Norris, 2003; Scarborough et al., 2019) can therefore not be confirmed on the basis of our data. In addition, people who trust traditional media more found three out of four types of statements "less" acceptable. This suggests a normative impact of traditional media on what is considered acceptable speech, which is an aspect that should also be investigated further.

(6) The results are less straightforward for experiences of discrimination and political attitudes. First, for the former, it seems that more frequent and diverse experiences instead led to a small de-sensitizing effect in terms of microaggression and hate speech. In contrast, experiences with gender-specific experiences contribute to a sensitizing effect for microaggressions. On a general level, this is consistent with previous findings that prior experiences with hurtful statements influence perceptions of future ones (e.g., Mummendey, 2019). In specific terms, however, the differentiated results also raise new questions concerning the further research that would be fruitful. Nevertheless, at this point, at least the importance of further research on microaggressions again becomes clear.

(7) Political attitudes were not as important as assumed. Two effects, however, point in the assumed direction. Political self-positioning proved

influential for microaggressions, with people on the right perceiving these statements as more acceptable, and the AfD preference has the same effect in terms of derogatory statements. These results are also consistent with previous studies on perceptions of freedom of expression (e.g., Costello et al., 2019; Petersen, 2021). However, while it has often been assumed that political factors, in particular, could be a key factor in these debates, we could not find such definite results within the framework of our method.

(8) Finally, we found surprising effects concerning the neutral control items, for which there could be methodological as well as substantial reasons: either the other overtly discriminating statements affected the evaluation of the neutral statements by producing a halo effect, or predispositions actually led to a differing assessment of even everyday statements. Further, the respective (imagined) context—i.e., the conversation partners and the situation—can also decisively influence the perception of the statements. As we cannot judge, either way, further research is needed here as well.

Although our results provide important indications of whether and how perceptions of statements vary across society, they are limited by some apparent factors. For example, the topic of statements about women is, of course, only one of many relevant to the question of social consensus. However, even within the topic itself, the number of statements examined was very limited with only two items per level. While they provided an initial basis for our exploratory approach, they did not, of course, cover the full range of discriminatory statements. Accordingly, the results cannot immediately be generalized to other topics or groups and thus require replications for other forms of discrimination and a wider range of statements. In addition, it should be mentioned that the number of respondents was relatively small due to the application of a split ballot design. Accordingly, besides the thematic limitation, the results are also limited with regard to sample size.

Overall, it can be summarized that the study provides initial insights into whether a social consensus exists in terms of which statements are currently socially accepted and which are considered unspeakable as well as what predispositions might impact these perceptions. Based on an initial exploratory framework that examined different types of discrimination, this study broadens the view from a specific phenomenon to a general perspective of freedom of expression. Accordingly, this conceptual view could also serve as a starting point for future research regarding the tension between freedom of expression and freedom from discrimination. While free speech is a fundamental basis of democracy, it can be simultaneously argued that democratic societies will do well to find common ground about

what they regard as discriminatory speech in order to strengthen respectful public discourse. As far as this study is concerned, the controversial area of conflict does not seem to evolve around offensive discrimination but rather more subtle expressions. This study thus provides initial insights on the kinds of potentially controversial statements on which there is a general consensus and those that we as a society still need to discuss. Future research should follow up on these conclusions.

References

Abi-Hassan, S. (2017). Populism and gender. In C. R. Kaltwasser, P. A. Taggart, P. O. Espejo, & P. Ostiguy (Eds.), *The Oxford handbook of populism* (pp. 426–444). Oxford University Press.

Berlin District Court. (2020). *Beschluss vom 21.01.2020, 27 AR 17/19: Anspruch auf Auskunft zu einem beleidigenden Nutzer eines sozialen Netzwerks* [Decision of 21.01.2020, 27 AR 17/19: Right to disclosure of information about an offending user of a social network]. https://gesetze.berlin.de/bsbe/document/KORE534882020

Boyle, K. (2001). Hate speech–- The United States versus the rest of the world? *Main Law Review*, *53*, 487–502. https://digitalcommons.mainelaw.maine.edu/mlr/vol53/iss2/7

Budziszewska, M., Hansen, K., & Bilewicz, M. (2014). Backlash over gender-fair language: The impact of feminine job titles on men's and women's perception of women *Journal of Language and Social Psychology*, *33*(6), 681–691. https://doi.org/10.1177/0261927x14544371

Chen, G. M., Pain, P., Chen, V. Y., Mekelburg, M., Springer, N., & Troger, F. (2020). 'You really have to have a thick skin.' A cross-cultural perspective on how online harassment influences female journalists. *Journalism*, *21*(7), 877–895. https://doi.org/10.1177/1464884918768500

Christley, O. (2021). Traditional gender attitudes, nativism, and support for the radical right. *Politics & Gender*, 1–27. https://doi.org/10.1017/S1743923X21000374

Costello, M., Hawdon, J., Bernatzky, C., & Mendes, K. (2019). Social group identity and perceptions of online hate. *Sociological Inquiry*, *89*(3), 427–452. https://doi.org/10.1111/soin.12274

Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V. M. (2010). Prejudice, stereotyping and discrimination: Theoretical and empirical overview. In J. F. Dovidio (Ed.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 3–28). SAGE. https://doi.org/10.4135/9781446200919.n1

Fawzi, N., Steindl, N., Obermaier, M., Prochazka, F., Arlt, D., Blöbaum, B., Dohle, M., Engelke, K. M., Hanitzsch, T., Jackob, N., Jakobs, I., Klawier, T., Post, S., Reinemann, C., Schweiger, W., & Ziegele, M. (2021). Concepts, causes and consequences of trust in news media – A literature review and framework. *Annals of the International Communication Association*, *45*, 154–174. https://doi.org/10.1080/23808985.2021.1960181

Foster, M. (2009). The dynamic nature of coping with gender discrimination: Appraisals, strategies and well-being over time. *Sex Roles*, *60*(9), 694–707. https://doi.org/10.1007/s11199-008-9568-2

Garton Ash, T. (2016). *Free speech: Ten principles for a connected world*. Yale University Press.

Gustafsson Sendén, M., Bäck, E. A., & Lindqvist, A. (2015). Introducing a gender-neutral pronoun in a natural gender language: The influence of time on attitudes and behavior. *Frontiers in Psychology, 6,* 893. https://doi.org/10.3389/fpsyg.2015.00893

Hong, M. (2020). Meinungsfreiheit und ihre Grenzen [Freedom of speech and its limits]. *Politik und Zeitgeschichte*. https://www.bpb.de/apuz/306444/meinungsfreiheit-und-ihre-grenzen

Inglehart, R., & Norris, P. (2003). *Rising tide: Gender equality and cultural change around the world.* Cambridge University Press.

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision Making, 8*(4), 407–424. http://dx.doi.org/10.2139/ssrn.2182588

Kister, K. (2021, September 15). Zwickau: Die absurde Entscheidung des Verwaltungsgerichts Chemnitz [Zwickau: The absurd decision of the Administrative Court of Chemnitz]. *Süddeutsche Zeitung*. https://www.sueddeutsche.de/meinung/verwaltungsgericht-chemnitz-neonazis-zwickau-1.5410637

Knight Foundation. (2019). *Freedom of speech at college campuses*. https://knightfoundation.org/reports/free-expression-college-campuses/

Köcher, R. (2019). *Grenzen der freiheit* [Limits to freedom]. Allensbach Institute for Public Opinion Research. https://www.ifd-allensbach.de/fileadmin/user_upload/FAZ_Mai2019_Meinungsfreiheit.pdf

Leets, L. (2002). Experiencing hate speech. Perceptions and responses to anti-Semitism and antigay speech. *Journal of Social Issues, 58*, 341–361. http://dx.doi.org/10.1111/1540-4560.00264

Lilienfeld, S. O. (2017). Microaggressions: Strong claims, inadequate evidence. *Perspectives on Psychological Science*, *12*(1), 138–169. https://doi.org/10.1177%2F1745691616659391

Love, R. (2021). Swearing in informal spoken English: 1990s–2010s. *Text & Talk*, *41*(5-6), 739–762. https://doi.org/10.1515/text-2020-0051

Lui, P. P., & Quezada, L. (2019). Associations between microaggression and adjustment outcomes: A meta-analytic and narrative review. *Psychological Bulletin*, *145*(1), 45–78. https://doi.org/10.1037/bul0000172

Marker, K. (2013). Know your enemy. Zur Funktionalität der Hassrede für wehrhafte Demokratien [On the functionality of hate speech for defensible democracies]. In J. Meibauer (Ed.), *Hassrede. Interdisziplinäre Beiträge zu einer aktuellen Diskussion* [Hate speech. Interdisciplinary contributions to a current discussion] (pp. 59–94). Gießener Elektronische Bibliothek. http://geb.uni-giessen.de/geb/volltexte/2013/9251/pdf/HassredeMeibauer_2013.pdf

Mummendey, A., Kessler, T., & Otten, S. (2009). Sozialpsychologische Determinanten–- Gruppenzugehörigkeit und soziale Kategorisierung [Social psychological determinants–- Group membership and social categorization.]. In A. Beelmann, K. J. Jonas, A. Beelmann, & K. J. Jonas (Eds.), *Diskriminierung und Toleranz. Psychologische Grundlagen und Anwendungsperspektiven* [Discrimination and Tolerance. Psychological foundations and application perspectives] (pp. 43–60). Wiesbaden.

Noelle-Neumann, E. (1996). *Öffentliche Meinung: Die Entdeckung der Schweigespirale* [Public opinion: The discovery of the spiral of silence]. Ullstein.

Petersen, T. (2013). *Tatsächliche und gefühlte Intoleranz* [Actual and perceived intolerance]. Allensbach Institute for Public Opinion Research. https://www.ifd-allensbach.de/fileadmin/kurzberichte_dokumentationen/Maerz13_Intoleranz.pdf

Petersen, T. (2021). *Die Mehrheit fühlt sich gegängelt* [The majority feels it's being restrained]. Allensbach Institute for Public Opinion Research. https://www.ifd-allensbach.de/fileadmin/kurzberichte_dokumentationen/FAZ_Juni2021_Meinungsfreiheit.pdf

Pew Research Center. (2019). *A changing world: Global views on diversity, gender equality, family life and the importance of religion*. https://www.pewresearch.org/global/2019/04/22/how-people-around-the-world-view-gender-equality-in-their-countries/

Quiring, O., Jackob, N., Schemer, C., Jakobs, I., & Ziegele, M. (2020). „Das wird man doch noch sagen dürfen…" – Wahrgenommene Sprechverbote und ihre Korrelate ["One is still allowed to say that..." – Perceived speech prohibitions and their correlates]. In N. Jackob, O. Quiring, & M. Maurer (Eds.), *Traditionen und transformationen des öffentlichen: Soziale wahrnehmung und soziale kontrolle in zeiten gesellschaftlichen Wandels* [Traditions and transformations of the public sphere: Social perception and social control in times of social Change] (pp. 49–71). Springer VS. https://doi.org/10.1007/978-3-658-29321-5_3

Reimer, F. (2019). Freiheit der Meinungsäußerung: eine rechtsvergleichende perspektive – Deutschland [Freedom of expression: A comparative law perspective - Germany]. *EPRS | Wissenschaftlicher Dienst des Europäischen Parlaments*. https://www.europarl.europa.eu/RegData/etudes/STUD/2019/642269/EPRS_STU(2019)642269_DE.pdf

Scarborough, W. J., Sin, R., & Risman, B. (2019). Attitudes and the stalled gender revolution: Egalitarianism, traditionalism, and ambivalence from 1977 through 2016. *Gender & Society*, *33*(2), 173–200. https://doi.org/10.1177/0891243218809604

Strömbäck, J., Tsfati, Y., Boomgaarden, H. G., Damstra, A., Lindgren, E., Vliegenthart, R., & Lindholm, T. (2020). News media trust and its impact on media use: Toward a framework for future research. *Annals of the International Communication Association*, *44*, 139–156. https://doi.org/10.1080/23808985.2020.1755338

Struth, A. K. (2019). *Hassrede und freiheit der meinungsäußerung* [Hate speech and freedom of speech]. Springer.

Sue, D. W. (2010). *Microaggressions in everyday life: Race, gender, and sexual orientation*. John Wiley & Sons.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *The American Psychologist*, *62*(4), 271-286. https://doi.org/10.1037/0003-066X.62.4.271

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Torino, G. C., Rivera, D. P., Capodilupo, C. M., Nadal, K. L., & Sue, D. W. (Eds.). (2018). *Microaggression theory: Influence and implications*. John Wiley & Sons.

Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, *63*(2), 221–243. https://doi.org/10.1111/jcom.12024

Wegner, J., Prommer, E., & Collado Seidel, C. (2020). Free speech under pressure: The effect of online harassment on literary writers. *Media and Communication*, *8*(4), 145–157. https://www.ssoar.info/ssoar/bitstream/document/70232/2/ssoar-mediacomm-2020-4-wegner_et_al-Free_Speech_Under_Pressure_The.pdf

Williams, D. R., & Mohammed, S. A. (2013). Racism and health I: Pathways and scientific evidence. *The American Behavioral Scientist*, *57*(8). https://doi.org/10.1177/0002764213487340

Wojcieszak, M., Sobkowicz, P., Yu, X., & Bulat, B. (2021). What information drives political polarization? Comparing the effects of in-group praise, out-group derogation, and evidence-based communications on polarization. *The International Journal of Press/Politics*. https://doi.org/10.1177%2F19401612211004418

Carsten Reinemann *(Dr. phil., U of Mainz, 2003) is a Professor at the Department of Media and Communication, LMU Munich, Germany. His research currently deals with populism, extremism, changing opinion power in hybrid media systems and issues of freedom of speech. Wolfram Peiser was a PostDoc at the U of Mainz when Carsten Reinemann graduated there in 1997 (M.A.). They became colleagues when Carsten Reinemann started his PhD in Mainz in late 1997 until Wolfram Peiser went to LMU Munich in 2006 to become a Professor. In 2008, their ways crossed again when Carsten Reinemann also joined the Munich department as a Professor.*

Anna-Luisa Sacher *is a Research Associate and PhD student at the Department of Media and Communication (IfKW) at Ludwig-Maximilians-Universität Munich, Germany. Her research interests focus on the perception, use and impact of political communication, where she is particularly concerned with aspects of the tension between freedom of speech and freedom from discrimination.*