



# Examining strategic diversity communication on social media using supervised machine learning: Development, validation and future research directions

Joep Hofhuis<sup>a,b,\*</sup>, João Gonçalves<sup>a</sup>, Pytrik Schafraad<sup>c,d</sup>, Biyao Wu<sup>c,e</sup>

<sup>a</sup> Department of Media & Communication, Erasmus University Rotterdam, the Netherlands

<sup>b</sup> Amsterdam School of International Business (AMSIB), Amsterdam University of Applied Sciences, the Netherlands

<sup>c</sup> Department of Communication Science, University of Amsterdam, the Netherlands

<sup>d</sup> Department of Communication and Cognition, Tilburg University, the Netherlands

<sup>e</sup> Department of Marketing, VU University Amsterdam, the Netherlands

## ARTICLE INFO

### Keywords:

Social media  
Strategic communication  
Diversity  
Diversity perspectives  
Computational methods  
Supervised machine-learning

## ABSTRACT

In this paper, we present a digital tool named *Diversity Perspectives in Social Media* (DivPSM) which conducts automated content analysis of strategic diversity communication in organizational social media posts, using supervised machine-learning. DivPSM is trained to identify whether a post makes mention of diversity or a diversity-related issue, and to subsequently code for the presence of three diversity dimensions (cultural/ethnic/racial, gender, and LHGBTQ+ diversity) and three diversity perspectives (the moral, market, and innovation perspectives). In Study 1, we describe the training and validation of the instrument, and examine how it performs compared to human coders. Our findings confirm that DivPSM is sufficiently reliable for use in future research. In study 2, we illustrate the type of data that DivPSM generates, by analyzing the prevalence of strategic diversity communication in social media posts ( $n = 84,561$ ) of large organizations in the Netherlands. Our results show that in this context gender diversity is most prevalent, followed by LHGBTQ+ and cultural/ethnic/racial diversity. Furthermore, gender diversity is often associated with the innovation perspective, whereas LHGBTQ+ diversity is more often associated with the moral perspective. Cultural/ethnic/racial diversity does not show strong associations with any of the perspectives. Theoretical implications and directions for future research are discussed at the end of the paper.

## 1. Introduction

Management of diversity is considered to be increasingly important for modern organizations (Galinsky et al., 2015; Plaut, 2010). Previous research shows that successful workplace inclusion of individuals with different cultural backgrounds, genders, sexual orientations, etc. can have a profound impact on equality in society, and can simultaneously provide benefits for organizational productivity (Homan, 2019). Most of us have seen organizations express their commitment to workplace diversity on company websites, and in social media posts, job advertisements, and periodicals such as annual reports. The connection between diversity and strategic communication has been well established (Bear et al., 2010; Hon & Brunner, 2000), and is further emphasized in more recent work that suggests that diversity communication may be an important driver of social change (Christensen et al., 2013; Schoeneborn

et al., 2020).

Until recently, academic inquiry into the topic of strategic diversity communication was limited to manual content analyses of diversity statements (e.g. Jonsen et al., 2021; Maier & Ravazzani, 2019; Maior-escu-Murphy, 2022) or effect studies using experimental designs (e.g. Hofhuis, Van der Zee, et al., 2016; Jansen et al., 2021; Windscheid et al., 2017). However, recent developments in computational social science, especially in the area of automated content analysis, have provided new opportunities for studying organizational communication on a much larger scale (Van Zoonen & Van der Meer, 2016). New avenues of exploration have opened up, which may advance our understanding of how and why organizations communicate about diversity, how these messages evolve over time, and how they impact different stakeholder groups (Hofhuis, Schafraad, et al., 2023; Lazer et al., 2020; Zhou, 2021).

In this paper, we present a digital tool that enables scholars to take

\* Correspondence to: Erasmus University Rotterdam, ESHCC Department of Media & Communication, PO Box 1738, 3000 DR Rotterdam, the Netherlands.

E-mail address: [j.hofhuis@eshcc.eur.nl](mailto:j.hofhuis@eshcc.eur.nl) (J. Hofhuis).

<https://doi.org/10.1016/j.pubrev.2024.102431>

Received 28 April 2023; Received in revised form 8 November 2023; Accepted 31 January 2024

Available online 10 February 2024

0363-8111/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

advantage of these opportunities through automated coding of strategic diversity communication in organizational social media posts, using supervised machine learning. The tool is trained to recognize which dimensions of diversity (racial/ethnic/cultural, gender, LHGBTQ+) are mentioned in each post. Furthermore, to examine the post content on a deeper level, it also codes which underlying rationale for diversity management the organization is communicating. For this purpose, we make use of seminal work by Ely and Thomas (2001), who identified three motivations for organizations to formulate and implement diversity policy, termed *diversity perspectives*. The new tool that is presented in this paper, named *Diversity Perspectives in Social Media* (DivPSM), distinguishes between a *moral perspective* (e.g., 'We want to eliminate discrimination and work towards an inclusive society'), *market perspective* (e.g. 'A diverse workforce makes our organization better able to adapt to a complex market'), and *innovation perspective* (e.g. 'Workplace diversity is associated with greater problem solving capacity, flexibility and innovation'). DivPSM is able to reliably identify which of these diversity perspectives are mentioned in individual social media posts, and in relation to which diversity dimension.

In this paper, we will first describe the development and validation of DivPSM (Study 1). Next, we provide an illustration of how DivPSM can be used, by examining the prevalence of diversity dimensions and perspectives in Twitter and Facebook posts of a sample of large organizations in the Netherlands (Study 2). In the final section of this paper, we discuss implications and future research directions.

## 2. Theoretical Background

### 2.1. Strategic diversity communication

A growing body of research focuses on how topics related to workplace diversity are represented in strategic organizational communication, and how these messages relate to organizational and societal outcomes. Some scholars have interpreted diversity activities as a form of Corporate Social Responsibility (CSR) policy (e.g. Maier & Ravazzani, 2019; Starostka-Patyk et al., 2015). Diversity management has indeed been revealed as one of the CSR activities that is most valued by the market (Bird et al., 2007). Bear and colleagues (2010), for example, show that the percentage of women on a company's executive board has a positive impact on CSR ratings, which in turn positively affects organizational reputation. A recent content analysis of corporate diversity communication on Twitter also confirmed that consistent CSR communication on diversity policies received positive feedback from stakeholders (Mairescu-Murphy, 2022).

Diversity communication has also been identified as a form of Corporate Social Advocacy (CSA), through which organizations showcase their support for social movements and civic engagement (Zhang, 2022; Zhou, 2021). For example, a number of large multinational corporations have come together to support gay marriage legislations (Wettstein & Baur, 2016), while others publicly supported the Black Lives Matter movement (Ciszek & Logan, 2018). Although such communications can be met with mixed reactions, diversity-related CSA activities are becoming increasingly common, and may form a substantial proportion of strategic diversity communication (Hofhuis, Schafraad, et al., 2023; Rim et al., 2020).

Finally, diversity communication can also be approached from the perspective of employer branding (Edwards & Kelan, 2011). By broadcasting their diversity policy and the rationale behind it, organizations aim to improve their image and enhance their reputation as an attractive employer, enhancing their position on the job market, and increasing their ability to hire promising talent (Fields et al., 2005; Hofhuis et al., 2015; Jonsen et al., 2021). This approach has further advantages, because successful diversity communication may not only have an impact on external stakeholders and potential employees, but also reaches internal stakeholders, including current employees. As such, it may influence perceived inclusion, diversity climate, and job satisfaction of

different groups within the organization (Christensen et al., 2013; Hofhuis et al., 2012; Hofhuis et al., 2016).

Previous studies have analyzed the presence of diversity statements in different types of communication channels, and in different national contexts. Much prior work, including some of the earliest studies (e.g. Point & Singh, 2003), focuses on diversity communication on corporate websites. This body of research has revealed an increasing prevalence of diversity statements among organizations in many different countries, such as the U.K. (Guerrier & Wilson, 2011), U.S. (Abulbasal, 2023; Rim et al., 2020), France, Germany, Spain (Jonsen et al., 2021), and Portugal (Barbosa & Cabral-Cardoso, 2010). Other scholars have performed content analyses of different communication channels, such as annual reports (Hofhuis, et al., 2023; Zhou, 2021), job market communication (Casper et al., 2013; Gaucher et al., 2011), corporate media (Kroon et al., 2018) and management communication (Bader et al., 2019; Ng & Sears, 2020) in a variety of different regions and sectors.

### 2.2. Diversity communication on social media

In recent years, we have seen an increase in strategic diversity communication on social media channels, with Twitter (e.g. Mairescu-Murphy, 2022) and Facebook (e.g. Zhang, 2022) being most prominent. Social media have become instrumental to the public-facing identity of organizations, and allow them to respond to trends or current affairs in a much more dynamic way than was possible with more traditional channels, such as annual reports and websites (Kent & Li, 2020). Therefore, examining diversity statements in social media posts enables us to better understand which types of diversity communication are seen as relevant by organizations, and how communication professionals are responding to current events.

The growth of strategic social media use by organizations has provided an abundance of data that reflects diversity-related activities and perspectives. However, much of this data remains unexamined. Until recently, large-scale content analyses of social media posts tended to be unfeasible due to the time-consuming and costly realities of human coding methods. The first aim of the present research is to provide scholars with a new tool that enables examination of large datasets in a much more efficient and cost-effective way, through automated coding.

### 2.3. Diversity dimensions

Another limitation of the existing literature on strategic diversity communication is that, while some comparative research exist (e.g. Jonsen et al., 2021), the majority of studies focused on a single diversity dimension, such as culture (Hofhuis, Schafraad, et al., 2023), gender (Windscheid et al., 2018), LHGBTQ+ (Zhou, 2021), or age (Kroon et al., 2018). However, in a world of complex identities and increasing focus on intersectionality, it is important that scholars are able to examine the combination of, and interplay between, different diversity dimensions (Bowleg & Bauer, 2016; Shore et al., 2009). To facilitate such work, the second aim of the digital tool that is presented in this paper is to identify the prevalence of different dimensions in social media communication.

The first dimension that is included in this study is diversity in cultural, ethnic, or racial background. It is important to note that culture, ethnicity and race are different constructs that have widely diverging definitions, social implications, and academic interpretations (Chao & Otsuki-Clutter, 2011; Phinney & Ong, 2007; Vietze et al., 2023). However, in organizational communication, we see that the use of these constructs is highly dependent on the societal context of the organization. For example, Western European organizations are more likely to focus on cultural/ethnic diversity, whereas US organizations tend to focus more on racial diversity (Jonsen et al., 2021). To accommodate the use of our tool across different national and cultural boundaries, we have therefore decided to aggregate these specific labels into one diversity dimension, termed 'cultural/ethnic/racial diversity'.

Second, our new tool will be able to capture whether social media

posts contain reference to gender diversity, defined as the mention of diversity along traditional binary gender boundaries, such as the presence of men and women in different position in the organizations. Third, the tool captures references to LHGBTQ+ diversity which is defined as the mention of sexual orientations, or any form of non-binary or fluid gender/sex in the organizational context.

It is important to be aware of the many other diversity dimensions that may be included in organizational communications, such as age, physical/mental ability, religion, language, etc. However, in the current iteration of our tool, it was decided to only focus on the three dimensions defined above, because they currently appear to be the most prevalent in organizational discourse (see also Jonsen et al., 2021; Shore et al., 2009). Future iterations of the tool may be trained to also recognize other dimensions, and compare their prevalence and co-occurrence.

#### 2.4. Diversity perspectives

There is growing recognition that the success of diversity communication is a function of not only the simple mentioning of diversity, or even the concrete policies and activities that organizations promote. In addition, the public has become more interested in the underlying motivations and rationale for organizations to focus on enhancing diversity, termed *diversity perspectives*.<sup>1</sup> (Dwertmann et al., 2016; Leslie & Flynn, 2022). The third aim of the tool presented in this paper is to reliably identify which diversity perspectives are being mentioned in organizational social media posts.

The most widely used framework for examining diversity perspectives was developed by Thomas and Ely (1996), and verified in a qualitative study among a sample of American organizations (Ely & Thomas, 2001). They identified three specific perspectives, which organizations may adopt in their drive towards diversification. Firstly, the *Moral Perspective* reflects the idea that enhancing diversity in the workplace can be seen as an ethical or moral obligation that organizations have towards society. By promoting diversity, an organization implies it is a socially responsible institution, providing equal opportunities to all cultural groups and aiming to reduce discrimination. The motivations for formulating diversity policy within this perspective may include the normative idea of 'doing good', external legal or regulatory incentives, or the aim of generating a positive impact on the organization's reputation (Bear et al., 2010; Bird et al., 2007; Podsiadlowski et al., 2013).

Secondly, the *Market Perspective*, refers to the notion that most organizations operate in a society or market which is inherently culturally diverse. Hence, having a diverse workforce is a valuable tool for gaining knowledge about, and access to, different groups of stakeholders. An example is a supermarket in a culturally diverse neighborhood, which matches the cultural backgrounds of its employees with those of its customers to provide the best customer service. The same principle can be applied to other types of organizations; a governmental organization must have a diverse workforce in order to understand and meet the needs of different groups within the society it serves; a health care provider must be able to communicate effectively about health-related issues in terms that all groups in society are able to comprehend (Ashikali & Groeneveld, 2015; Hofhuis et al., 2018).

Thirdly, the *Innovation Perspective* reflects the idea that cultural diversity may have direct benefits for internal processes within the organization, such as increased problem solving potential, creativity, flexibility, and innovation. Research in the area of information elaboration and creativity in teams has provided evidence that diversity may indeed lead organizations to become more flexible and innovative in completing their assigned tasks (Van Knippenberg et al., 2004). When

<sup>1</sup> In the literature, the terms diversity strategy, orientation, approach, frame, and perspective are often used synonymously. In line with Ely and Thomas (2001), whose work forms the basis of the current research, we employ the latter.

the opportunity is given to voice different viewpoints, the presence of deviant opinions may increase creative thinking (De Dreu & West, 2001; Zhan, 2023). Previous studies suggest that under the right conditions, cultural diversity may increase the effectiveness of idea generation, knowledge sharing, and ultimately team performance (Hofhuis, Van der Rijt, et al., 2016; Nakui et al., 2011). In sum, organizations that employ this perspective believe that cultural diversity may lead to higher flexibility, creativity, and innovation, which in turn could increase the effectiveness of the organization as a whole (Rabl et al., 2020; Shipton et al., 2005).

#### 2.5. Automated content analysis of diversity perspectives

The tool that is presented in this paper makes use of supervised machine-learning to identify the presence of the moral, market, and innovation perspectives in organizational social media posts. The starting point for this work was the previous development of a similar tool (DivPAR; Hofhuis, Schafraad, et al., 2023), which is able to identify these perspectives in regard to cultural diversity in corporate annual reports, using a top-down dictionary based approach. Annual reports contain large amounts of text, and organizations are increasingly likely to use them to outline their diversity perspectives and initiatives in great detail (Stanton & Stanton, 2002). As such, for the previous tool it was possible to develop a set of specific coding rules, based on the presence of certain word combinations or phrases in the text, making a dictionary-based approach to automated content analysis most suitable for this task (Boumans & Trilling, 2016; Hofhuis, Schafraad, et al., 2023). DivPAR has been shown to be a reliable and more efficient alternative to manual coding for this particular purpose. Since its inception, it has been adopted by scholars to investigate current trends in strategic diversity communication of, among others, large organizations in the Netherlands (Hofhuis, Santos, & Schafraad, 2023), and IT companies in the United States (Abulbasal, 2023).

However, to expand this research towards coding of social media posts, the previously used dictionary-approach was less feasible. Since social media posts are limited in length, authors resort to more subtle, varied, and indirect ways to communicate about diversity perspectives. In many cases, the underlying rationale is implied, rather than explicitly stated. Automated coding of diversity dimensions and perspectives in such communications requires a more flexible approach, which more closely resembles human interpretation. As such, we chose to use supervised machine-learning for this task.

Study 1 describes the development, training, and validation procedure of the new tool (DivPSM), and outlines how it performs compared to human coders. After the validation procedure was completed, we used DivPSM to code a sample of Tweets and Facebook posts by large Dutch organizations, to illustrate the type of data that the tool generates, and how it can be used to analyze trends in strategic diversity communication. The results of this coding process are described in study 2. Finally, we outline a number of potential future research directions, which we hope may inspire scholars to make use of DivPSM in their own work.

### 3. Study 1: development and validation of DivPSM

#### 3.1. DivPSM instrument design

Study 1 describes the development and validation of DivPSM (*Diversity Perspectives in Social Media*), which was designed to extract the following variables from organizational social media posts: 1) whether it mentions diversity or any diversity-related topic, whether it mentions 2) cultural/ethnic/racial diversity, 3) gender diversity, or 4) LHGBTQ+ diversity, and whether it mentions the 5) moral perspective, 6) market perspective, or 7) innovation perspective.

As explained above, due to the dynamic, subtle and often contextual nature of organizational communications on social media, a dictionary-

based approach was not feasible for this purpose. Instead, the task requires a more flexible coding procedure, that more closely resembles human interpretation (Albaugh et al., 2014). Therefore, we decided to design our tool using supervised machine-learning. We developed and trained a deep learning model based on Bidirectional Encoder Representations from Transformers (BERT; Devlin et al., 2019; Tenney et al., 2019). BERT has achieved state of the art or near state-of-the-art performance in many text classification tasks (González-Carvajal & Garrido-Merchán, 2021) and is widely available and used, which enhances its reusability by other researchers.

### 3.2. Sample

The sample for our validation study consist of Tweets and Facebook posts of large organizations in the Netherlands. The organizations were selected from *De Volkskrant Top 100* largest employers in the country (De Volkskrant, 2020). It is important to note that, although all of these organizations are based in the Netherlands, we only selected those who communicate in English on their social media channels. The majority of these organizations are multinationals, and have a substantial presence outside the Netherlands. The final sample included 54 organizations, from a wide variety of sectors, such as finance (e.g. ING, ABN AMRO), services (e.g. KPMG, E&Y), industry (e.g. Shell, Unilever), and non-profit (e.g. Greenpeace).

Social Media posts ( $n = 84,561$ ) were collected using a Twitter API and Facebook's Crowdtangle platform. The majority ( $n = 74,561$ ) were posts by the official Twitter accounts of the companies in the sample. In September 2021, Twitter's Standard API V1 was used to collect post by these accounts going back in time, upto the maximum allowed limit of 3200 per account. This dataset spanned a time period from 2010–2021.

The data were then supplemented with Facebook posts by the same companies' official accounts ( $n = 10,000$ ) that were collected using the Crowdtangle platform. We used Crowdtangle's filtering functionality to obtain English-language posts only. Public posts in the official Facebook company pages, including texts accompanying video or image posts, were scraped spanning a timeframe between 2009–2021. To comply with the platform's 10,000 scraping limit, we used Crowdtangle's default ranking of posts by number of interactions (likes, shares, comments, and reactions) and obtained the 10,000 with the most interactions.

### 3.3. Training and validation

#### 3.3.1. Step 1 – developing the manual codebook

The development and validation procedure for DivPSM consisted of 4 steps. In step 1, we developed a codebook for manual content analysis of the variables mentioned above, based on previous research that included similar coding tasks (e.g. Ely & Thomas, 2001; Hofhuis et al., 2015; Hofhuis, Schafraad, et al., 2023). We created two separate subsamples of social media posts: Subsample 1 ( $n = 200$ ) consisted of posts that were randomly chosen from the larger dataset, subsample 2 ( $n = 100$ ) consisted of posts that were identified by the researchers to contain text markers of diversity-related content, such as 'diversity', 'inclusion', 'ethnic', 'racial', 'multicultural', or 'gender equality'. Both subsamples were coded by a team of five human coders (the authors and one trained research assistant), who identified all seven variables under investigation. It is important to note that social media-posts by organizations also contain non-official language such as slang or emojis, which we accounted for in the coding instructions. To enhance reliability, results were compared and discussed among the coders, until agreement was reached on the operationalization of each variable. For example, discussion took place on when a post contains gender diversity vs. LHGBQ+ diversity, leading to the decision that any mention of male and/or female employees is considered gender diversity, whereas mention of other forms or combinations of gender are considered LHGBQ+ diversity. Furthermore, it was decided that a diversity

perspective is only coded as present when there is specific mention of a motivation, or rationale behind a diversity statement that is made in a post. Appendix A provides examples of the final operationalization and corresponding coding instructions for each variable in the codebook, as well as examples of posts that were found to contain the respective content.

#### 3.3.2. Step 2 – Validating the manual codebook

In step 2, the manual codebook was validated by examining inter-coder reliability. Another set of subsamples was selected: Subsample 3 ( $n = 400$ ) contained randomly chosen posts from the larger dataset, and subsample 4 ( $n = 200$ ) consisted of posts that were identified to contain the abovementioned diversity markers. The same five individuals manually coded these subsamples, and Krippendorff's Alpha (Hayes & Krippendorff, 2007) was calculated. For variables 1 (Presence of diversity statement), 2 (cultural/racial/ethnic), 3 (gender), and 4 (LHGBQ+), intercoder reliability was very high (Kripp. Alpha  $> .89$ ). For variables 5 (moral perspective) and 7 (innovation perspective), intercoder reliability was good (Kripp. Alpha = .83; .78). For variable 6 (market perspective), intercoder reliability was initially low (Kripp. Alpha = .58). After consulting among the team of coders, it was concluded that one of the manual coders did not identify a number of more subtle occurrences of this perspective (i.e. mention of diversity to connect to the community, or to serve clients). These codings were corrected in the final dataset, leading to sufficient reliability (Kripp. Alpha = .76). The corrected data were subsequently used in the following steps.

#### 3.3.3. Step 3 – training the algorithms

In step 3, the combined manual coding was used to train the first version of the machine-learning algorithms. The data was fed into a pretrained uncased multilingual BERT model using the *ktrain* python package (Maiya, 2022). *ktrain* was selected because of its simplicity and intuitive nature, enhancing the potential for widespread use of the trained models. In total, 5 models were trained, one for overall presence of diversity, one for each of the three diversity perspectives, and a multiclass model for the presence of the three diversity dimensions.

A particular challenge for the training process was the underrepresentation of some of the classes in our sample. Using a first version of the models, we further identified a subsample of 450 posts which were identified to be more likely to contain one of the three diversity perspectives. These posts were also manually coded, and the results were subsequently used to fine-tune the algorithm. Even with an initial sample of nearly 85,000 tweets and Facebook posts, and with the additional coding described above, positive examples for the market and innovation perspectives were insufficiently prevalent to reach acceptable levels of performance. Three measures were taken to minimize these issues: data augmentation, balancing of datasets, and learning rate tweaking.

Regarding data augmentation, for the three perspectives, the lead author created 6 examples for the market and innovation perspectives that could be seen as ideal types of that diversity perspective. The moral perspective did not require augmentation as it had sufficient positive cases. These were then fed into ChatGPT to generate 18 additional examples per perspective, curated by the research team to ensure that they still reflected the perspective. Because the transformers model underlying GPT is different from BERT, we trusted that this would help increase the generalization potential of our samples. These augmented examples, including the ideal types, were then included in the training data, but excluded from the test data to ensure that they would not artificially inflate reliability estimates.

While data augmentation increased the number of positive cases in underrepresented classes, there was still a substantial class imbalance in some classes. For instance, the ratio of positive to negative cases in the market perspective, within posts flagged for diversity, was 1:16. This resulted in the bias term of the algorithm overpowering any variance

during the training stage, and no positive cases were produced in the inference stage. To avoid this, models for the diversity perspectives were trained and validated in balanced datasets, where negative cases were omitted until there was an approximately equal number of positives and negatives. However, this also resulted in training set sizes that ranged from 126 (market perspective) to 1422 (overall diversity presence). Thus, to prevent high learning rates that would overstep global minima during gradient descent, triangular learning rates were employed, adjusting the maximum ceiling for the learning rate between 2e-5 and 2e-7, with smaller learning rates used for smaller datasets (Howard & Ruder, 2018). While balancing of the datasets may skew predictions to overestimate the presence of the perspectives or dimensions, confusion matrices during validation did not reveal any substantial bias towards false positives or false negatives.

3.3.4. Step 4 – validating the algorithms

In step 4 we trained and validated the performance of the algorithms, using all previously labelled examples (n = 1358), partitioned into 80/10/10 (training, validation and testing, respectively) for variable 1 and (70/20/10) for the other variables. Non-diversity examples were not used to train, validate and test diversity perspectives and dimensions. For each variable the partition was random with the following conditions: 1) augmented examples were not used in test sets; 2) negative examples were randomly removed from all variables except for diversity to balance the datasets.

We trained for a maximum of 250 epochs and used early stopping and loaded the weights from the best epoch according to validation loss. In practice these 250 epochs were never reached, as the model began to overfit before that point. Validation loss was used to assess the point to which the model began to overfit, we used 10% of the labelled dataset to validate the classifier for the presence of diversity, and 20% of the labelled dataset for the remaining variables. Given that these datasets were smaller, a larger percentage was required to ensure that validation was robust. Finally, to ensure that the models were not overfitting to validation set, a separate test set with 10% of the sample was used to confirm reliability estimates.

Accuracy and macro-F1 values for all variables are presented in Table 1. There is still room for improvement through retraining with larger datasets, particularly for the diversity perspectives. However, based on the results, we believe that the current version of DivPSM is sufficiently reliable to be used for inference.

4. Study 2: diversity dimensions and perspectives in social media posts by organizations in the Netherlands

4.1. Study aim and methods

Having described the development and validation of the instrument, the aim of study 2 is to illustrate how DivPSM can be used to generate datasets on the prevalence of strategic diversity communication in organizational social media posts. Specifically, we examined which forms of strategic diversity communication can be recognized in the full sample (n = 84,561) of Twitter and Facebook posts, which was described in Study 1. We used the final validated version of DivPSM to

code the full dataset, identifying the presence of different diversity dimensions and perspectives and how these correlate with each other.

4.2. Results

The first step in analyzing the prevalence of diversity statements in our sample of social media posts, was to examine DivPSM’s coding of variable 1, which indicates whether a post mentions diversity or any diversity-related topic. Our findings show that this was the case for 4.6% (n = 3880) of the posts in the sample. Further analyses were subsequently conducted on these posts only. Table 2. describes the outcomes of the automated coding process, and correlations between the variables under study.

The second step in our analyses was to examine the prevalence of the three different diversity dimensions. DivPSM coding showed that 26.8% (n = 1041) of diversity-related posts mentioned cultural/ethnic/racial diversity, 41.6% (n = 1616) mentioned gender diversity, and 36.2% (n = 1404) mentioned LHGBTQ+ diversity. It appears that in the current sample, gender diversity is the most commonly occurring diversity dimension, followed by LHGBTQ+ diversity, and cultural/ethnic/racial diversity. When interpreting bivariate correlations between the variables, an interesting pattern emerges. We see a strong negative correlation (r = -.41) between gender and LHGBTQ+ diversity, meaning that posts that contain one of those are much less likely to contain the other. No substantial correlations are observed with cultural/ethnic/racial diversity, meaning that this dimension does co-occur with the other dimensions, but not more than would be expected based on chance.

The third step in our analysis was to examine the prevalence of the three diversity perspectives. DivPSM coding showed that 22% (n = 852)

Table 2

Prevalence of and correlations between diversity dimensions and diversity perspectives in diversity-related social media posts (n = 3880) of Dutch organizations (n = 54).

Variable	Prevalence	r*					
		2	3	4	5	6	7
<i>Diversity Dimensions</i>							
2. Ethnic/Cultural/Racial Diversity	1041 (26.8%)	-	-0.07	0.00	0.06	0.06	0.08
3. Gender Diversity	1616 (41.6%)	-0.07	-	-0.41	-0.33	0.03	0.31
4. LHGBTQ+ Diversity	1404 (36.2%)	0.00	-0.41	-	0.48	0.04	-0.36
<i>Diversity Perspectives</i>							
5. Moral Perspective	852 (22.0%)	0.06	-0.33	0.48	-	0.16	-0.25
6. Market Perspective	137 (3.5%)	0.06	0.03	0.04	0.16	-	0.08
7. Innovation Perspective	1222 (31.5%)	0.08	0.31	-0.36	-0.25	0.08	-

\* due to the large sample size, all correlations stronger than .01 are statistically significant

Table 1

Accuracy and F1 values per variable.

Variable	n	Training Accuracy	Validation Accuracy	Validation F1	Test Accuracy	Test F1
1. Diversity	1358	97.5%	91.0%	.91	83.1%	.83
2. Cultural/racial/ethnic	131	93.9%	84.0%	.84	84.6%	.80
3. Gender	167	93.9%	78.2%	.78	82.4%	.80
4. LHGBTQ+	96	95.8%	88.9%	.88	80.0%	.80
5. Moral	272	85.6%	65.3%	.65	66.6%	.61
6. Market	124	72.7%	66.7%	.66	83.3%	.80
7. Innovation	106	93.7%	70.0%	.69	68.2%	.63

mentioned the moral perspective, 3.5% ( $n = 137$ ) mentioned the market perspective, and 31.5% ( $n = 1222$ ) mentioned the innovation perspective. It appears that the innovation perspective is most prevalent in diversity-related posts, followed by the moral perspective, which is also mentioned in a substantial number of cases. However, the market perspective appears quite rare in this context. When examining correlations, it is revealed that posts that contain the moral perspective are also more likely ( $r = .16$ ) to contain the market perspective, but less likely ( $r = -.25$ ) to contain the innovation perspective. This suggests that organizations may perceive the moral and innovation perspectives as incompatible, at least in communication on social media.

Finally, our dataset also allows us to examine the interplay between dimensions and perspectives. For cultural/ethnic/racial diversity, no noteworthy correlations are observed with any of the diversity perspectives. However, it appears that post that mention gender diversity are less likely ( $r = -.33$ ) to contain the moral perspective, and more likely ( $r = .31$ ) to contain the innovation perspective. For post that mention LHGBTQ+ diversity, an opposite pattern is observed; these are much more likely ( $r = .48$ ) to contain the moral perspective, and much less likely to contain the innovation perspective ( $r = -.36$ ). The relevance of these findings will be interpreted below.

## 5. Discussion and future research agenda

In study 1, we have described the development and validation of DivPSM, a new digital tool that allows scholars to conduct automatic content analysis of strategic diversity communication in organizational social media posts. The current version of DivPSM is sufficiently reliable for further use, but reliability may still be increased in future iterations. The latest version of DivPSM is open source, and will be made available to scholars on GitHub (<https://github.com/joephohhuis/divpsm>). Study 2 provides an illustration of how the tool can be used to analyze the prevalence and interplay between these variables in social media posts of a sample of Dutch organizations. Below, we will describe some of the theoretical implications of our work, and we provide a number of directions for future research, that we hope may inspire scholars to use DivPSM to advance knowledge on the topics of workplace diversity and strategic communication.

### 5.1. Overview of findings and theoretical implications

Firstly, our work may contribute to the understanding of strategic diversity communication on a theoretical level. The large-scale investigation of diversity communication may provide new insights in the role that organizations may play in both initiating as well as responding to social change. Recent theoretical advances have been made in the conceptualization of *formative CSR communication*, which reflects the idea that communication and practices are mutually and continuously constitutive, and thus influence each other across time (Schoeneborn et al., 2020). In line with this notion, this project is founded on the assumption that aside from diversity practices influencing communication, successful diversity communication may also be a driver for new diversity management practices and activities.

Previous work that has focused on the prevalence of diversity perspectives in organizational communication has been limited in scope, for example by focusing on only a single diversity dimension (Hofhuis, et al., 2023; Zhou, 2021), or by aggregating different diversity perspectives into broader categories (Dwertmann et al., 2016; Jansen et al., 2021). Using DivPSM, we are able to examine the interplay between different dimensions and perspectives, and identify whether some types of diversity statements are more likely to be used conjunctively. The preliminary results presented in Study 2 provide some empirical evidence that this may indeed be the case: among our sample of social media posts by Dutch organizations, gender diversity is often associated with the innovation perspective, whereas LHGBTQ+ diversity is more often associated with the moral perspective. This suggests that gender

diversity is more commonly perceived to have a positive impact on organizational performance, for example through enhanced creativity and problem solving ability, whereas LHGBTQ+ diversity is more commonly communicated as a form of Corporate Social Responsibility or Corporate Social Advocacy. This is an interesting finding by itself, because it implies that it is seen as more acceptable to use a 'business case' argumentation for promoting gender diversity, based on the notion that there may be inherent differences between the men and women, and that these differences can be used to enhance productivity. These types of assumptions have been called into question by scholars who argue that such a perspective may also have a negative impact on workplace inclusion of certain stigmatized groups (Georgeac & Rattan, 2022; Windscheid et al., 2017).

Conversely, LHGBTQ+ diversity communication tends to make more frequent use of the moral perspective, which could imply that individuals who feel connected to the LHGBTQ+ community may not be perceived as inherently different in their work behavior, but instead deserve to be treated fairly and should feel included in the organization on the basis of moral grounds (e.g. as social advocacy; Zhou, 2021). Based on the present study, we cannot test these assumptions outright, but we hope future scholars may be able to use DivPSM to further tease out the nuances of how organizations communicate about these different diversity dimensions and how this relates to actual workplace outcomes.

Furthermore, our findings shed new light on the way communication professionals in these organizations might rationalize the diversity policies of their respective employer, and on their expectations regarding public response to diversity statements. It also has important implications for the study of skepticism towards diversity communication, and the occurrence of pinkwashing, colorwashing, or other forms of non-authentic messages about diversity-related issues (e.g. Ma & Bentley, 2022), since their effects may also be dependent on the dimension under investigation.

What is also interesting to note is that in the sample described above, the innovation perspective is by far the most prevalent, followed by the moral perspective. The market perspective only occurs very rarely. This deviates significantly from previous work that reports cultural diversity perspectives in annual reports of a similar sample of Dutch organizations (Hofhuis, Schafraad, et al., 2023), in which the moral perspective was found to be the most common, followed by the market and innovation perspective. The fact that examining different communication channels leads to different findings, implies that organizations may not always be consistent in their diversity communication, and that different motivations or rationales may be present among different communication agents, or when targeting different audiences. This underlines the importance of studying these issues conjunctively, and examining the decision-making-process around diversity management, and how (and by whom) it is communicated with both internal and external stakeholders.

### 5.2. Future Research Directions

Attention for diversity in strategic organizational communication relates to contextual factors on micro, meso and macro levels. The instrument that we developed and validated in this study can be used to reliably investigate such relationships on a large scale. On the micro level, we can investigate the effectiveness of diversity statements as a public relations tool, by examining their effects on message engagement, individual perceptions of organizational reputation, and stakeholder reactions (cf. Maiorescu-Murphy, 2022; Zhang, 2022). On the meso-level, diversity perspectives can be related to organizational reputation, innovation, profitability, or diversity management initiatives (cf. Jansen et al., 2021). Here, we may also be able to uncover the extent to which organizations engage in pinkwashing, colorwashing, or any other form of diversity communication that may be perceived as inauthentic by the public (cf. Forehand & Grier, 2003; Hofhuis, Santos,

& Schafraad, 2023; Ma & Bentley, 2022). On the macro-level, our results can be linked to societal indicators or events such as economic crises or political developments (cf. Hofhuis, 2022; Korteweg & Triadafilopoulos, 2015; Meer et al., 2015). We can also conduct cross-national comparative studies on the prevalence of different types of strategic diversity communication, and how they relate to cultural values (i.e. World Value Survey), or national attitudes towards diversity (Jonsen et al., 2021; Zhou, 2021).

Finally, all of these effects can be studied longitudinally, taking into account developments over time and investigating current trends (cf. Hofhuis, 2022; Hofhuis, Schafraad, et al., 2023). Each of these investigations would add significant parts to the puzzle of understanding patterns in the prevalence of diversity perspectives in strategic communication, and how they relate to organizational and societal outcomes.


### 5.3. Further development of DivPSM

While DivPSM is already in a usable state for analysis, its accuracy and reliability can still be improved. One of the limitations of the current iteration of the tool is that it was trained on social media posts of organizations that are active in the Dutch economic context. While many of these organizations are also active internationally, we hope that future versions of the tool can be improved by training with larger and more diverse types of social media data, such as those from different regions around the globe, or by incorporating different social media platforms, such as LinkedIn or Mastodon.

Additionally, the BERT uncased model currently does not tokenize emojis, which may be an important part of diversity discourses on social media. As such, training on a model that is capable of processing these tokens would potentially increase accuracy.

Another important limitation of the current version of DivPSM, is that it is only capable of reliably identifying three diversity dimensions; cultural/ethnic/racial, gender, and LHGBTQ+. For further development of the tool, we intend to train it to also recognize other important dimensions that may be common in organizational social media posts, such as age diversity, religious diversity, neurodiversity, or inclusion of employees with disabilities. To do so would require manual coding and validation of a larger sample of posts, to ensure enough occurrences of these dimensions to be able to successfully train a machine-learning

## Appendix A. : Example operational definitions and example posts

Codebook item	Coder instructions	Example posts
Diversity	The concept may include any mention of diversity in any of its dimensions, or of diversity management practices and policies in the organizational context. The dimensions of diversity may include gender, religious beliefs, race, marital status, ethnicity, parental status, age, education, physical and mental ability, income, sexual orientation, occupation, language, geographic location, etc. The coders should be sensitive to expressions such as cultural, ethnicity, racial, national, multinational, diverse, variety, equality, equity, etc.	<p>“We are proud to support @CreativeEquals ‘Accelerate’ programme recognising* rising talent from multicultural/BAME backgrounds &amp; supporting them on the path to #Leadership. #DiversityandInclusion #Diversity #DiversityMatters.”</p> <p>“For our employees, the community &amp; world, we’re committed to driving gender equality &amp; diversity. #CatalystForChange”</p> <p>“Full support for our LGBTI+ employees on #ComingOutDay. An internal broadcast tomorrow will share Juliana Fochi’s Silveira Araujo story with our colleagues: “I’m glad to feel accepted for who I am, though I know not everyone feels the same yet.</p> 
Moral perspective	The moral perspective reflects the idea that enhancing diversity in the workplace can be seen as an ethical or moral obligation that organizations have toward society. By promoting diversity, an organization implies it is a socially responsible institution, aiming to reduce discrimination and enhance equality, equity and inclusion. The coders should be sensitive to terms such as moral, ethical, fair, equal, discrimination, racism, moral responsibility, social responsibility, etc.	<p>#AkzoNobel #Inclusion”</p> <p>“Creating a diverse and inclusive organisation requires intentional advocacy to combat racism and elimination of unconscious bias,” says our CEO @AlanJope Which is why we’re joining @WEF’s Racial Justice in Business initiative.”</p> <p>“Racial issues are human rights issues. No such thing as neutral in racism. #BlackLivesMatter”</p> <p>“New research uncovered that ethnic minorities don’t feel represented in tech advertising. The #ChangeTheFace Alliance aims to change this and support #diverse and #inclusive workforces in the #technology industry through resources, research, tools and tips.”</p>

(continued on next page)

model.

Finally, it is important to note that at present, DivPSM is only capable of coding English-language posts. Considering the status of English as a lingua franca within many multinational organizations, it was logical to focus on training these language models first. However, in the future, we hope to be able to expand DivPSM to include other language models, which would allow for much more accurate cross-national comparisons, and open up possibilities of studying how strategic diversity communications affects different groups of local and regional stakeholders.

### 5.4. Conclusion

Above, we have described the development and validation of a new digital tool for automatic content analysis of strategic diversity communication in organizational social media posts (DivPSM). We have provided a first illustration of how the instrument can be used, by showing the interplay between diversity dimensions and diversity perspectives in posts by large organizations in the Netherlands. Finally, we have provided a number of theoretical implications and future research directions, that we hope may inspire others to use our instrument in their own work. We believe DivPSM to be a valuable addition to the toolbox of diversity science, and expect that its results will add to our understanding of strategic diversity communication, how it relates to individual, organizational and societal outcomes, and how it may ultimately contribute to equality and inclusion in the workplace.

### Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Declaration of Competing Interest

None.

### Acknowledgements

The authors thank Veerle van Oosterom for her contributions to data collection and coding, and Jana Vietze for her feedback on earlier drafts of the manuscript.

(continued)

Codebook item	Coder instructions	Example posts
Market perspective	The market perspective refers to the notion that most organizations operate in a society or market which is inherently diverse. Hence, having a diverse workforce is a valuable tool for gaining knowledge about, and access to, different groups of stakeholders. The coders should be sensitive to terms such as market, stakeholders, groups, communities, society, partners, serve, reflect, mirror, connect, understand.	<p>“We have asked our partners to make two promises: one on living wage/income &amp; one on supplier equity, diversity &amp; inclusion. Why? Because we believe that purpose-led partnerships can drive growth, while also doing good for people and the planet.”</p> <p>“Inclusive leadership helps companies achieve their purpose, foster growth and build long-term value.”</p> <p>“Did you know that the UN considers cultural diversity an important driver of development, innovation and economic growth? We believe it’s one of our greatest strengths. That’s why we strive to create a workplace where we celebrate our differences. #WorldDayForCulturalDiversity”</p>
Innovation perspective	The innovation perspective reflects the idea that cultural diversity may have direct benefits for internal processes within the organization, such as increased learning potential and innovation.” Coders should be sensitive to terms such as innovation, improve performance, competitive advantage, competitive edge, creativity.	<p>“Speak up, speak out. Event organiser Bizzabo found that 69% of event speakers were male. Diverse panels can reflect society and bring new ideas to the table. #IWD2020”</p> <p>“Are you happier and more productive in a welcoming and inclusive environment? So is Niels den Heeten, a Senior Innovation Project Manager who also leads our True Colors LGBTI+ inclusion initiative</p>
Cultural/Ethnic/Racial diversity	Cultural/ethnic/racial diversity refers to issues concerning the presence and/or position of cultural, ethnic, or racial minorities in the organization.	<p>. Learn more here: <a href="https://t.co/5bkXN0x0t8#MeetAkzoNobel#EmployeeStories">https://t.co/5bkXN0x0t8#MeetAkzoNobel#EmployeeStories</a>”</p> <p>“Did you know that the UN considers cultural diversity an important driver of development, innovation and economic growth? We believe it’s one of our greatest strengths. That’s why we strive to create a workplace where we celebrate our differences. #WorldDayForCulturalDiversity”</p> <p>“Creating a diverse and inclusive organisation requires intentional advocacy to combat racism and elimination of unconscious bias,” says our CEO @AlanJope Which is why we’re joining @WEF’s Racial Justice in Business initiative.”</p>
Gender diversity	Gender diversity refers to issues concerning the presence and position of men and women, or male/female employees in the organization.	<p>“Womenpower empowers and fuels innovation. Happy #InternationalWomensDay from #BMW”</p> <p>“Still a huge issue in the finance sector for women across all parts of the industry. #genderpaygap #equalpay”</p>
LHGTBQ+ diversity	LHGTBQ+ diversity refers to issues concerning the position of Lesbian, Homosexual, Gay, Transsexual, Bisexual, Queer, or other (e.g. non-binary, gender fluid) types of sexual identity or orientations in the organization.	<p>Happy international coming out day</p> <p>Be who you are and don't stop until you are proud!                  @accor @mgalleryhotels                  #MGalleryMoments #accorhotels #beproud #StoriesThatStay [...] #mgalleryhotels #thenetherlands #travel #photooftheday #experience #stories #INKredible”</p> <p>“Happy International Coming Out Day! Across LexisNexis we recognize the great work our Pride, PROUD, sexual orientation and gender identity Employee Resource Groups are doing to foster a culture of inclusion, safety and respect for our LGBTQ+ colleagues around the world. Discover their stories: <a href="https://bit.ly/43o8GVz">https://bit.ly/43o8GVz</a> #ComingOutDay #LNDiversity”</p>

References

Abulbasal, R. (2023). Diversity in tech: Analyzing the mismatch between corporate framing and impact [PhD Dissertation]. Utah State University.

Albaugh, Q., Soroka, S., Joly, J., Loewen, P., Sevenans, J., & Walgrave, S. (2014). Comparing and combining machine learning and dictionary-based approaches to topic coding. *Proceedings of the 7th Annual Comparative Agendas Project (CAP) Conference*, 12–14.

Ashikali, T., & Groeneveld, S. (2015). Diversity management in public organizations and its effect on employees’ affective commitment: The role of transformational leadership and the inclusiveness of the organizational culture. *Review of Public Personnel Administration*, 35(2), 146–168. <https://doi.org/10.1177/0734371x13511088>

Bader, A. K., Kemper, L. E., & Froese, F. J. (2019). Who promotes a value-in-diversity perspective? a fuzzy set analysis of executives’ individual and organizational characteristics. *Human Resource Management*, 58(2), 203–217. <https://doi.org/10.1002/hrm.21946>

Barbosa, I., & Cabral-Cardoso, C. (2010). Equality and diversity rhetoric: one size fits all? globalization and the Portuguese context. *Equality, Diversity and Inclusion: An International Journal*, 29(1), 97–112. <https://doi.org/10.1108/02610151011019237>

Bear, S., Rahman, N., & Post, C. (2010). The impact of board diversity and gender composition on corporate social responsibility and firm reputation. *Journal of Business Ethics*, 97(2), 207–221. <https://doi.org/10.1007/s10551-010-0505-2>

Bird, R., Hall, A. D., Momente, F., & Reggiani, F. (2007). What corporate social responsibility activities are valued by the market? *Journal of Business Ethics*, 76(2), 189–206. <https://doi.org/10.1007/s10551-006-9268-1>

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, 4(1), 8–23. <https://doi.org/10.1080/21670811.2015.1096598>

Bowleg, L., & Bauer, G. (2016). Invited reflection: quantifying intersectionality. *Psychology of Women Quarterly*, 40(3), 337–341. <https://doi.org/10.1177/03616843166654282>

Casper, W. J., Wayne, J. H., & Manegold, J. G. (2013). Who will we recruit? targeting deep-and surface-level diversity with human resource policy advertising. *Human Resource Management*, 52(3), 311–332. <https://doi.org/10.1002/hrm.21530>

Chao, R. K., & Otsuki-Clutter, M. (2011). Racial and ethnic differences: sociocultural and contextual explanations. *Journal of Research on Adolescence*, 21(1), 47–60. <https://doi.org/10.1111/j.1532-7795.2010.00714.x>

Christensen, L. T., Morsing, M., & Thyssen, O. (2013). CSR as aspirational talk. *Organization*, 20(3), 372–393. <https://doi.org/10.1177/1350508413478310>

Ciszek, E., & Logan, N. (2018). Challenging the dialogic promise: how Ben & Jerry’s support for black lives matter fosters dissensus on social media. *Journal of Public Relations Research*, 30(3), 115–127. <https://doi.org/10.1080/1062726X.2018.1498342>

De Dreu, C. K. W., & West, M. A. (2001). Minority dissent and team innovation: the importance of participation in decision making. *Journal of Applied Psychology*, 86(6), 1191–1201. <https://doi.org/10.1037/0021-9010.86.6.1191>



- De Volkskrant. (2020, February 15). De top 100 grootste werkgevers in Nederland. De Volkskrant. www.volkskrant.nl.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. <https://doi.org/10.48550/arXiv.1810.04805>
- Dwertmann, D. J., Nishii, L. H., & Van Knippenberg, D. (2016). Disentangling the fairness & discrimination and synergy perspectives on diversity climate: Moving the field forward. *Journal of Management*, 42(5), 1136–1168. <https://doi.org/10.1177/0149206316630380>
- Edwards, M. R., & Kelan, E. K. (2011). Employer Branding and Diversity: Foes or Friends? In *Branded Lives. The Production and Consumption of Meaning at Work*. Edward Elgar Publishing.
- Ely, R. J., & Thomas, D. A. (2001). Cultural diversity at work: the effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, 46(2), 229–273. <https://doi.org/10.2307/2667087>
- Fields, D. L., Goodman, J. S., & Blum, T. C. (2005). Human resource dependence and organizational demography: a study of minority employment in private sector companies. *Journal of Management*, 31(2), 167–185.
- Forehand, M. R., & Grier, S. (2003). When is honesty the best policy? the effect of stated company intent on consumer skepticism. *Journal of Consumer Psychology*, 13(3), 349–356. [https://doi.org/10.1207/S15327663JCP1303\\_15](https://doi.org/10.1207/S15327663JCP1303_15)
- Galinsky, A. D., Todd, A. R., Homan, A. C., Phillips, K. W., Apfelbaum, E. P., Sasaki, S. J., Richeson, J. A., Olayon, J. B., & Maddux, W. W. (2015). Maximizing the gains and minimizing the pains of diversity: a policy perspective. *Perspectives on Psychological Science*, 10(6), 742–748. <https://doi.org/10.1177/1745691615598513>
- Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1), 109–128. <https://doi.org/10.1037/a0022530>
- Georgeac, O. A. M., & Rattan, A. (2022). The business case for diversity backfires: detrimental effects of organizations' instrumental diversity rhetoric for underrepresented group members' sense of belonging. *Journal of Personality and Social Psychology*, 124(1), 69–108. <https://doi.org/10.1037/pspi0000394>
- González-Carvajal, S., & Garrido-Merchán, E. C. (2021). *Comparing BERT against traditional machine learning text classification* (arXiv:2005.13012). arXiv. <https://doi.org/10.48550/arXiv.2005.13012>
- Guerrier, Y., & Wilson, C. (2011). Representing diversity on UK company web sites. *Equality, Diversity and Inclusion: An International Journal*, 30(3), 183–195. <https://doi.org/10.1108/02610151111124932>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89.
- Hofhuis, J. (2022). Comparing cultural diversity perspectives among public service employees in the Netherlands in 2008 and 2018. *Equality, Diversity and Inclusion: An International Journal*, 41(5), 726–738. <https://doi.org/10.1108/EDI-01-2021-0002>
- Hofhuis, J., Mensen, M., ten Den, L. M., van den Berg, A. M., Koopman-Draijer, M., van Tilburg, M. C., Smits, C. H. M., & de Vries, S. (2018). Does functional diversity increase effectiveness of community care teams? The moderating role of shared vision, interaction frequency, and team reflexivity. *Journal of Applied Social Psychology*, 48(10), 535–548. <https://doi.org/10.1111/jasp.12533>
- Hofhuis, J., Santos, G., & Schafraad, P. (2023). *Diversiteitskampioen of Colorwashing? De samenhang tussen diversiteitscommunicatie en diversiteitsuitkomsten in het Nederlandse bedrijfsleven*. Erasmus University Rotterdam, Erasmus Research Centre for Media, Communication and Culture (ERMeCC).
- Hofhuis, J., Schafraad, P., Trilling, D., Luca, N., & Van Manen, B. (2023). Automated content analysis of cultural diversity perspectives in annual reports (DivPAR): Development, validation, and future research agenda. *Cultural Diversity and Ethnic Minority Psychology*, 29(1), 74–84. <https://doi.org/10.1037/cdp0000413>
- Hofhuis, J., Van der Rijt, P. G. A., & Vlug, M. (2016). Diversity climate enhances work outcomes through trust and openness in workgroup communication. *SpringerPlus*, 5(1), 1–14. <https://doi.org/10.1186/s40064-016-2499-4>
- Hofhuis, J., Van der Zee, K. I., & Otten, S. (2012). Social identity patterns in culturally diverse organizations: the role of diversity climate. *Journal of Applied Social Psychology*, 42(4), 964–989. <https://doi.org/10.1111/j.1559-1816.2011.00848.x>
- Hofhuis, J., Van der Zee, K. I., & Otten, S. (2015). Measuring employee perception on the effects of cultural diversity at work: Development of the Benefits and Threats of Diversity Scale. *Quality & Quantity*, 49(1), 177–201. <https://doi.org/10.1007/s11135-013-9981-7>
- Hofhuis, J., Van der Zee, K. I., & Otten, S. (2016). Dealing with differences: the impact of perceived diversity outcomes on selection and assessment of minority candidates. *The International Journal of Human Resource Management*, 27(12), 1319–1339. <https://doi.org/10.1080/09585192.2015.1072100>
- Homan, A. C. (2019). Dealing with diversity in workgroups: preventing problems and promoting potential. *Social and Personality Psychology Compass*, 13(5), Article e12465. <https://doi.org/10.1111/spc3.12465>
- Hon, L. C., & Brunner, B. (2000). Diversity issues and public relations. *Journal of Public Relations Research*, 12(4), 309–340. [https://doi.org/10.1207/S1532754XJPRR1204\\_2](https://doi.org/10.1207/S1532754XJPRR1204_2)
- Howard, J., & Ruder, S. (2018). Universal Language Model Fine-tuning for Text Classification (arXiv:1801.06146). arXiv. <https://arxiv.org/abs/1801.06146>
- Jansen, W. S., Kröger, C., Van der Toorn, J., & Ellemers, N. (2021). The right thing to do or the smart thing to do? How communicating moral or business motives for diversity affects the employment image of Dutch public and private sector organizations. *Journal of Applied Social Psychology*, 51(7), 746–759. <https://doi.org/10.1111/jasp.12783>
- Jonsen, K., Point, S., Kelan, E. K., & Griebler, A. (2021). Diversity and inclusion branding: A five-country comparison of corporate websites. *The International Journal of Human Resource Management*, 32(3), 616–649. <https://doi.org/10.1080/09585192.2018.1496125>
- Kent, M. L., & Li, C. (2020). Toward a normative social media theory for public relations. *Public Relations Review*, 46(1), Article 101857. <https://doi.org/10.1016/j.pubrev.2019.101857>
- Kortweg, A. C., & Triadafilopoulos, T. (2015). Is multiculturalism dead? Groups, governments and the 'real work of integration. *Ethnic and Racial Studies*, 38(5), 663–680. <https://doi.org/10.1080/01419870.2014.907925>
- Kroon, A. C., Van Selm, M., Ter Hoeven, C. L., & Vliegenthart, R. (2018). Reliable and unproductive? Stereotypes of older employees in corporate and news media. *Ageing & Society*, 38(1), 166–191.
- Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margets, H., Nelson, A., Salganik, M. J., Strohmaier, M., Vespignani, A., & Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, 369(6507), 1060–1062. <https://doi.org/10.1126/science.aaz8170>
- Leslie, L. M., & Flynn, E. (2022). Diversity ideologies, beliefs, and climates: A review, integration, and set of recommendations. *Journal of Management*, 1–28. <https://doi.org/10.1177/01492063221086238>
- Ma, L., & Bentley, J. M. (2022). Can strategic message framing mitigate the negative effects of skeptical comments against corporate-social-responsibility communication on social networking sites? *Public Relations Review*, 48(4), Article 102222. <https://doi.org/10.1016/j.pubrev.2022.102222>
- Maier, C. D., & Ravazzani, S. (2019). Bridging diversity management and CSR in online external communication. *Corporate Communications*, 24(2), 269–286. <https://doi.org/10.1108/CCLJ-01-2018-0015>
- Maioreescu-Murphy, R. D. (2022). Business-centered versus socially responsible corporate diversity communication. An assessment of stakeholder (dis)agreement on Twitter. *Public Relations Review*, 48(1), Article 102138. <https://doi.org/10.1016/j.pubrev.2021.102138>
- Maiya, A. S. (2022). *ktrain: A Low-Code Library for Augmented Machine Learning* (arXiv:2004.10703). arXiv. <https://doi.org/10.48550/arXiv.2004.10703>
- Meer, N., Mouritsen, P., Faas, D., & de Witte, N. (2015). Examining 'Postmulticultural' and Civic Turns in the Netherlands, Britain, Germany, and Denmark. *American Behavioral Scientist*, 59(6), 702–726. <https://doi.org/10.1177/0002764214566496>
- Nakui, T., Paulus, P. B., & Van der Zee, K. I. (2011). The role of attitudes in reactions toward diversity in workgroups. *Journal of Applied Social Psychology*, 41(10), 2327–2351. <https://doi.org/10.1111/j.1559-1816.2011.00818.x>
- Ng, E. S., & Sears, G. J. (2020). Walking the Talk on Diversity: CEO Beliefs, Moral Values, and the Implementation of Workplace Diversity Practices. *Journal of Business Ethics*, 164(3), 437–450. <https://doi.org/10.1007/s10551-018-4051-7>
- Phinney, J. S., & Ong, A. D. (2007). Conceptualization and measurement of ethnic identity: Current status and future directions. *Journal of Counseling Psychology*, 54(3), 271–281. <https://doi.org/10.1037/0022-0167.54.3.271>
- Plaut, V. C. (2010). Diversity science: Why and how difference makes a difference. *Psychological Inquiry*, 21(2), 77–99. <https://doi.org/10.1080/10478401003676501>
- Podsiadlowski, A., Gröschke, D., Kogler, M., Springer, C., & Van der Zee, K. I. (2013). Managing a culturally diverse workforce: Diversity perspectives in organizations. *International Journal of Intercultural Relations*, 37(2), 159–175. <https://doi.org/10.1016/j.ijintrel.2012.09.001>
- Point, S., & Singh, V. (2003). Defining and dimensionalising diversity: Evidence from corporate websites across Europe. *European Management Journal*, 21(6), 750–761.
- Rabl, T., del Carmen Triana, M., Byun, S.-Y., & Bosch, L. (2020). Diversity management efforts as an ethical responsibility: How employees' perceptions of an organizational integration and learning approach to diversity affect employee behavior. *Journal of Business Ethics*, 161(3), 531–550. <https://doi.org/10.1007/s10551-018-3849-7>
- Rim, H., Lee, Y., & Yoo, S. (2020). Polarized public opinion responding to corporate social advocacy: Social network analysis of boycotters and advocates. *Public Relations Review*, 46(2), Article 101869. <https://doi.org/10.1016/j.pubrev.2019.101869>
- Schoeneborn, D., Morsing, M., & Crane, A. (2020). Formative perspectives on the relation between CSR communication and CSR practices: pathways for walking, talking, and T(w)alking. *Business & Society*, 59(1), 5–33. <https://doi.org/10.1177/0007650319845091>
- Shipton, H., Fay, D., West, M., Patterson, M., & Birdi, K. (2005). Managing People to Promote Innovation. *Creativity and Innovation Management*, 14(2), 118–128.
- Shore, L. M., Chung-Herrera, B. G., Dean, M. A., Ehrhart, K. H., Jung, D. I., Randel, A. E., & Singh, G. (2009). Diversity in organizations: Where are we now and where are we going? *Human Resource Management Review*, 19(2), 117–133. <https://doi.org/10.1016/j.hrmr.2008.10.004>
- Stanton, P., & Stanton, J. (2002). Corporate annual reports: Research perspectives used. *Accounting, Auditing & Accountability Journal*, 15(4), 478–500. <https://doi.org/10.1108/09513570210440568>
- Starostka-Patyk, M., Tomski, P., & Zawada, M. (2015). Diversity management as a part of corporate social responsibility policy. *Procedia Computer Science*, 65, 1038–1045. <https://doi.org/10.1016/j.procs.2015.09.055>
- Tenney, I., Das, D., & Pavlick, E. (2019). *BERT Rediscovered the Classical NLP Pipeline* (arXiv:1905.05950). arXiv. <https://doi.org/10.48550/arXiv.1905.05950>
- Thomas, D. A., & Ely, R. J. (1996). Making differences matter. *Harvard Business Review*, 1–12.
- Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, 89(6), 1008–1022. <https://doi.org/10.1037/0021-9010.89.6.1008>
- Van Zoonen, W., & Van der Meer, T., G. L. A. (2016). Social media research: The application of supervised machine learning in organizational communication

- research. *Computers in Human Behavior*, 63, 132–141. <https://doi.org/10.1016/j.chb.2016.05.028>
- Vietze, J., Schwarzenthal, M., Moffitt, U., & Civitillo, S. (2023). Beyond ‘migrant background’: How to select relevant, social justice oriented, and feasible social categories in educational research. *European Journal of Psychology of Education*, 38 (1), 389–408. <https://doi.org/10.1007/s10212-022-00611-2>
- Wettstein, F., & Baur, D. (2016). Why should we care about marriage equality?": Political advocacy as a part of corporate responsibility. *Journal of Business Ethics*, 138(2), 199–213. <https://doi.org/10.1007/s10551-015-2631-3>
- Windscheid, L., Bowes-Sperry, L., Jonsen, K., & Morner, M. (2018). Managing organizational gender diversity images: A content analysis of german corporate websites. *Journal of Business Ethics*, 152(4), 997–1013. <https://doi.org/10.1007/s10551-016-3292-6>
- Windscheid, L., Bowes-Sperry, L., Mazei, J., & Morner, M. (2017). The paradox of diversity initiatives: When organizational needs differ from employee preferences. *Journal of Business Ethics*, 145(1), 33–48. <https://doi.org/10.1007/s10551-015-2864-1>
- Zhan, M. M. (2023). Learning from the diverse perspectives and voice of newcomers: A contingency model. *Management Communication Quarterly*, 37(3), 423–450. <https://doi.org/10.1177/08933189221112048>
- Zhang, X. A. (2022). What constitutes great IDEA? An examination of corporate diversity communication on facebook and external and internal stakeholder reactions. *Public Relations Review*, 48(5), Article 102254. <https://doi.org/10.1016/j.pubrev.2022.102254>
- Zhou, A. (2021). Communicating corporate LGBTQ advocacy: A computational comparison of the global CSR discourse. *Public Relations Review*, 47(4), Article 102061. <https://doi.org/10.1016/j.pubrev.2021.102061>