

Gegründet im Auftrag der
Deutschen Gesellschaft für
Publizistik- und Kommuni-
kationswissenschaft e. V.
(DGPuK).

Established on behalf of
Deutsche Gesellschaft für
Publizistik- und Kommuni-
kationswissenschaft e. V.
(DGPuK).

Aus dem Inhalt

Full Paper

Elisabeth Schmidbauer & Jörg Haßler

Why do people (not) donate their data? A scoping review on factors influencing participation in data donation studies and strategies to enhance it

Tobias Schmidt, Jonas Rieger, Jörg Rahnenführer, Astrid Viciano & Holger Wormer

Beyond fact-checking: Harnessing AI for news quality assessment

Full Paper & Extended Abstract

Luna Laufer, Franziska Selzer, Julia Stockwald, Annika Schleithoff, Nathalie L. Neuendorf, Anna-Luisa Sacher & Carsten Reinemann

Freie Rede auf dem Campus? Einflüsse von politischer Kongruenz und Sanktionsängsten auf die Meinungsfreiheitswahrnehmung und Redebereitschaft deutscher Studierender

Full Paper

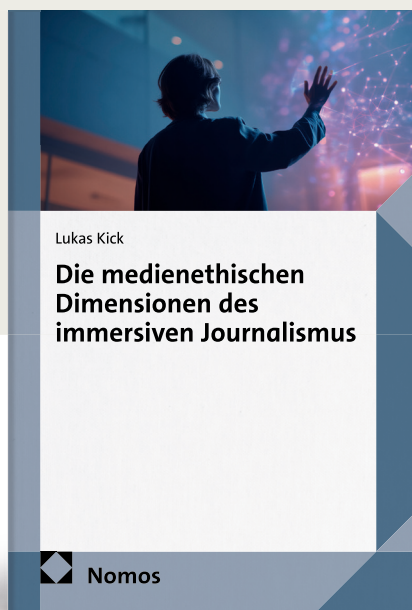
Claudia Thoms, Tobias Schrimpf & Andreas Reich

Climate Protest—What is it good for? Linking news coverage on German climate movements and reader comments

Elizabeth Prommer, Deb Verhoeven & Skadi Loist
Influential film producers and their networks: The formation of creative teams and gender inequality



Where Innovation meets Responsibility



Die medienethischen Dimensionen des immersiven Journalismus

By Dr. Lukas Kick

2026, 311 pp., pb., € 84.00

ISBN 978-3-7560-4032-2

E-Book 978-3-7489-7236-5

In German

Technical progress constantly creates new narrative possibilities for journalists and media producers. Immersive journalism is one of these technologies. It opens up additional news perspectives for the audience through a virtual extension of the perceived environment. This raises not only technical but also ethical challenges that media professionals

and audiences must address. Therefore, this book examines the media-ethics aspects resulting from the use of immersive journalism with its technical and narrative characteristics. Both theoretical and empirical perspectives are incorporated into the discussion.

Available digitally at [inlibra.com](https://www.inlibra.com)

Available in bookstores or free of shipping costs at [nomos-shop.de](https://www.nomos-shop.de)

Returns are at the addressee's risk and expense. All prices include VAT



Nomos

Gegründet im Auftrag der Deutschen Gesellschaft für Publizistik- und Kommunikationswissenschaft e. V. (DGPUK).

Established on behalf of Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e. V. (DGPUK).

Herausgeber/Publisher: Prof. Dr. Marko Bachl, Berlin | Prof. Dr. Christine Lohmeier, Salzburg | Prof. Dr. Helena Stehle, Münster | Asst. Prof. Dr. Anna Wagner, Nijmegen

Redaktion/Editorial Office:

Ruth Kasdorf, Hochschule Wismar, Fakultät Gestaltung, Philipp-Müller-Straße 14, 23966 Wismar

Internationaler Beirat/International Editorial Board: Prof. Dr. Jan van den Bulck, University of Michigan | Prof. Dr. Leopoldina Fortunati, Faculty of Education of the University of Udine, Italy | Dr. Beate Josephi, Edith Cowan University, Australia | Prof. Dr. Robin Mansell, London School of Economics and Political Science, UK | Prof. Dr. Dietram A. Scheufele, Department of Life Sciences Communication and College of Agricultural & Life Sciences, University of Wisconsin, USA | Prof. Dr. Peter J. Schulz, Institute of Communication and Health, University of Lugano, Switzerland | Prof. Dr. David Tewksbury, Department of Communication, University of Illinois at Urbana-Champaign, USA | Prof. Dr. Katerina Tsetsura, Gaylord College of Journalism and Mass Communication at the University of Oklahoma, USA | Prof. Dr. Gabriel Weimann, Department of Communication, University of Haifa, Israel

DGPK Beirat/DGPK Editorial Board: Mag. Dr. Ariadne Neureiter, FG Werbekommunikation | Dr. Julia Niemann-Lenz, FG Methoden der Publizistik- und Kommunikationswissenschaft | Jessica Kühn, M.A., FG Medienpädagogik | Jun.-Prof. Dr. Vera Katzenberger, FG Journalistik/Journalismusforschung | Dr. Erik Koenen, FG Kommunikationsgeschichte | Prof. Dr. Lara Nikola Wolfers, FG Rezeptions- und Wirkungsforschung | Jun.-Prof. Dr. Jessica Kunert, FG Mediensport und Sportkommunikation | Dr. Esther Greussing, FG Digitale Kommunikation | Dr. Katharina Christ, FG Mediensprache – Medien Diskurse | Dr. Philipp Müller, FG Kommunikation und Politik | Dr. Charmaine Voigt, FG Medien, Öffentlichkeit und Gesellschaft | Dr. Seraina Tarnutzer, FG Visuelle Kommunikation | Dr. Alexander Ort, FG Gesundheitskommunikation | Dr. Niels G. Mede, FG Wissenschaftskommunikation | Prof. Dr. Lars Rademacher, FG Kommunikations- und Medienethik | Dr. Anne Grüne, FG Internationale und Interkulturelle Kommunikation | Dr. Franziska Thiele, FG Soziologie der Medienkommunikation | Dr. Benno Viererbl, FG PR- und Organisationskommunikation | Prof. Dr. Britta Gossel, FG Medienökonomie

Aus dem Inhalt

FULL PAPER

Why do people (not) donate their data? A scoping review on factors influencing participation in data donation studies and strategies to enhance it

Warum spenden Menschen ihre Daten (nicht)? Ein Scoping-Review zu Faktoren, die die Teilnahme an Datenspende-Studien beeinflussen und Strategien zur Verbesserung dieser

Elisabeth Schmidbauer & Jörg Haßler 299

FULL PAPER

Beyond fact-checking: Harnessing AI for news quality assessment

Beyond Fact-Checking: KI-gestützte Qualitätsbewertung im Journalismus

Tobias Schmidt, Jonas Rieger, Jörg Rahnenführer, Astrid Viciano & Holger Wormer 332

FULL PAPER

Freie Rede auf dem Campus? Einflüsse von politischer Kongruenz und Sanktionsängsten auf die Meinungsfreiheitswahrnehmung und Redebereitschaft deutscher Studierender

Free speech on campus? The impact of political congruence and fear of sanctions on perceptions of freedom of speech and willingness to speak out among German university students

Luna Laufer, Franziska Selzer, Julia Stockwald, Annika Schleithoff, Nathalie L. Neuendorf, Anna-Luisa Sacher & Carsten Reinemann 359

EXTENDED ABSTRACT

Free speech on campus? The impact of political congruence and fear of sanctions on perceptions of freedom of speech and willingness to speak out among German university students

Freie Rede auf dem Campus? Einflüsse von politischer Kongruenz und Sanktionsängsten auf die Meinungsfreiheitswahrnehmung und Redebereitschaft deutscher Studierender

Luna Laufer, Franziska Selzer, Julia Stockwald, Annika Schleithoff,

Nathalie L. Neuendorf, Anna-Luisa Sacher & Carsten Reinemann 395

FULL PAPER

Climate Protest—What is it good for? Linking news coverage on German climate movements and reader comments

Klimaprotest – was nützt er wirklich? Eine Untersuchung von Artikeln über deutsche Klimabewegungen und dazugehörigen Leserkomentaren

Claudia Thoms, Tobias Schrimpf & Andreas Reich 402

FULL PAPER

Influential film producers and their networks: The formation of creative teams and gender inequality

Einflussreiche Filmproduzenten und ihre Netzwerke: Die Zusammensetzung von Filmteams und Geschlechterungleichheit

Elizabeth Prommer, Deb Verhoeven & Skadi Loist 430

Impressum

Herausgeber/Publisher: Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e. V. (DGPK), vertreten durch Prof. Dr. Marko Bachl, Berlin | Prof. Dr. Christine Lohmeier, Salzburg | Prof. Dr. Helena Stehle, Münster | Asst. Prof. Dr. Anna Wagner, Nijmegen

Redaktion/Editorial Office:

Ruth Kasdorf, Hochschule Wismar, Fakultät Gestaltung, Philipp-Müller-Straße 14, 23966 Wismar
e-mail: scm@nomos.de

Verlag/Publishing Company: Nomos Verlagsgesellschaft mbH & Co. KG,
Waldseestraße 3-5, 76530 Baden-Baden, Germany

Telefon/Phone: 0049 (0)7221/2104-0, Fax: 0049 (0)7221/2104-899

Internet: www.scm.nomos.de

Nachdruck und Vervielfältigung/Reprint and Reproduction: Die Zeitschrift und alle in ihr enthaltenen einzelnen Beiträge sind urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Der Nomos Verlag beachtet die Regeln des Börsenvereins des Deutschen Buchhandels e.V. zur Verwendung von Buchrezensionen.

Erscheinungsweise/Publication: viermal jährlich/four times a year

ISSN Online: 2192-4007

Hinweise für Autorinnen und Autoren finden Sie unter/Hints for authors: www.scm.nomos.de

FULL PAPER

Why do people (not) donate their data? A scoping review on factors influencing participation in data donation studies and strategies to enhance it

Warum spenden Menschen ihre Daten (nicht)? Ein Scoping-Review zu Faktoren, die die Teilnahme an Datenspende-Studien beeinflussen und Strategien zur Verbesserung dieser

Elisabeth Schmidbauer & Jörg Haßler

Elisabeth Schmidbauer (M. A.), LMU Munich, Department of Media and Communication, Akademiestraße 7, 80799 Munich, Germany. Contact: Elisabeth.Schmidbauer@ifkw.lmu.de. ORCID: <https://orcid.org/0009-0002-0343-1343>

Jörg Haßler (Assoc. Prof.), University of Oslo, Department of Media and Communication, Gaustadalléen 21, 0349 Oslo, Norway. Contact: joerg.hassler@media.uio.no. ORCID: <https://orcid.org/0000-0003-2907-5228>



Why do people (not) donate their data? A scoping review on factors influencing participation in data donation studies and strategies to enhance it

Warum spenden Menschen ihre Daten (nicht)? Ein Scoping-Review zu Faktoren, die die Teilnahme an Datenspende-Studien beeinflussen und Strategien zur Verbesserung dieser

Elisabeth Schmidbauer & Jörg Haßler

Abstract: Data donations have recently emerged as research method, partly enabled by legislation such as the GDPR. A relevant challenge they face are low participation rates, which can introduce bias and limit the potential of data donations to harness digital trace data for studying behavior in various contexts. We conducted an interdisciplinary scoping literature review to map research on participation in data donation studies, outlining factors that might discourage participation and strategies to address these issues. To ensure methodological clarity, data donations are defined by clear criteria: The donation is an active and unpaid, but at times incentivized, transfer of intangible, reproducible, and retrospective, trace data. Furthermore, data donations are voluntary and informed, and primarily motivated by contributing to the common good. Based on this definition, which also served as the inclusion criteria for our study, we analyzed 30 studies. Included literature predominantly focuses on health research and hypothetical scenarios, exploring attitudes rather than actual donations. Although the investigated literature is diverse, several factors are discussed to influence willingness to donate. These include respondent characteristics such as gender differences and technical skills, situational factors such as privacy concerns and motivations including egoism, study-related attributes including required data types and who receives the data, and societal-level factors such as contextual influences and culture. Strategies to increase willingness include general approaches like offering benefits and the right communication, as well as study design-related strategies, particularly cultivating trust, e.g., through credible collaborations and appropriate privacy protection.

Keywords: Data donations, participation-bias, literature review, trace data

Zusammenfassung: Datenspenden haben sich, begünstigt durch gesetzliche Rahmenbedingungen wie die DSGVO, als Forschungsmethode etabliert. Eine zentrale Herausforderung sind jedoch niedrige Teilnahmequoten, die zu Verzerrungen führen können und das Potenzial von digitalen Spurdaten zur Analyse von Verhalten einschränken. Mithilfe eines interdisziplinären Scoping Literature Reviews systematisieren wir in diesem Beitrag den Forschungsstand zur Teilnahme an Datenspendenstudien sowie Einflussfaktoren und mögliche Strategien zur Steigerung der Teilnahmebereitschaft. Wir definieren Datenspenden als aktiven, unbezahlten – aber mitunter incentivierten – Transfer von immateriellen reproduzier-

baren, rückblickenden Daten. Zudem sind Spenden freiwillig, informiert, mit Gemeinwohlbezug. Auf Basis dieser Definition, die zugleich als Einschlusskriterium diene, analysierten wir 30 Studien. Die Literatur stammt überwiegend aus der Gesundheitsforschung, fokussiert auf hypothetische Szenarien und untersucht primär Einstellungen statt tatsächlicher Teilnahme. Dennoch lassen sich zentrale Einflussfaktoren identifizieren, darunter individuelle Merkmale wie Geschlecht und technische Fähigkeiten, situative Aspekte wie Datenschutzbedenken oder Egoismus, studienspezifische Eigenschaften wie Datentypen und -Empfänger sowie gesellschaftliche Kontexte. Generelle Strategien zur Erhöhung der Teilnahmebereitschaft umfassen beispielsweise Anreize zu schaffen und zielgerichtete Kommunikation sowie den Aufbau von Vertrauen, etwa durch glaubwürdige Kooperationen und transparente Datenschutzmaßnahmen.

Schlagwörter: Datenspenden, Teilnahmeverzerrung, Literaturübersicht, Spurendaten

1. Introduction

A considerable proportion of daily activities now occurs within the digital sphere. Online behavior, whether intentional or not, leaves digital trace data, defined as “records of activity [...] undertaken through an online information system” (Howison et al., 2011, p. 769). These traces are rich data sources that enable more accurate and elaborate measurement of various phenomena (Araujo et al., 2017; Ohme et al., 2021; Scharkow, 2016; Wedel et al., 2024). They also help mitigate biases common in self-reported measures, such as reactivity (Baumgartner et al., 2023) or social desirability (Boeschoten et al., 2022).

Acquiring trace data is challenging because companies often withhold data and limit researchers’ access (Breuer et al., 2022, p. 2). One approach to improve accessibility is data donation, in which individuals contribute their personal data to research. A prominent example is the donation of data download packages (DDPs). Under the European Union’s General Data Protection Regulation (GDPR; EU Regulation 2016/679, 2016), users can request access to data platforms have stored about them (e.g., Pak et al., 2022, p. 135). Platforms provide this data in form of DDPs. These DDPs can then be shared with researchers, providing access to trace data on user behavior across platforms (Boeschoten et al., 2022; van Driel et al., 2022). For instance, Instagram DDPs may include data on likes, direct messages, and search queries (van Driel et al., 2022, p. 272).

Data gathered through donations can enhance research across disciplines, including communication and health, by improving data quality and providing insights into digital behavior. In communication studies, for example, Google DDP donations can be used to examine how citizens inform themselves about political topics before a referendum (Blassnig et al., 2023). In health research, data donations can support the measurement of physical activity and the analysis of related health outcomes (Keusch, Struminskaya, et al., 2024).

However, trace data, such as data donations, are subject to various biases and limitations (Amaya et al., 2020; Boeschoten et al., 2022; Hase & Haim, 2024). One stems from DDPs themselves. Because platforms create and control DDPs, their

structure and content can change at any time. This can introduce inconsistencies, ambiguities, and missing information, potentially compromising their reliability in research (Hase et al., 2024). To reduce errors in data donations, research emphasizes the need for best practices and clear recommendations (Carrière et al., 2025; Ohme & Araujo, 2022).

Much of the literature highlights data donations' susceptibility to non-response bias (Bietz et al., 2019; Ohme et al., 2023; Welbers et al., 2024). Participation rates are often low (Ohme et al., 2021; Pak et al., 2022), making it difficult to obtain representative samples (Ohme et al., 2023; Welbers et al., 2024). Improving understanding of the factors that predict non-response (e.g., Kmetty et al., 2024; Pfiffner & Friemel, 2023; Strycharz et al., 2024) is therefore essential for developing strategies to mitigate this bias.

This scoping review examines the literature on participation bias in data donation studies. It aims to identify respondent characteristics (e.g., age, gender), situational factors (i.e., motivations and study attributes), and societal influences shaping *why people do (not) donate their data*. It also seeks to identify strategies to increase participation.

2. Theory

2.1 What are data donations?

Data donation is a user-centric method of trace data collection in which participants share data with researchers (Boeschoten et al., 2022; Breuer et al., 2022; Haim & Hase, 2023; Ohme et al., 2023). As a distinct method, it must be differentiated from other forms of data sharing to understand the factors influencing participation. We therefore define data donations based on seven characteristics.

Data donations require participants' 1) *active involvement* in contributing data they hold (e.g., Bietz et al., 2019; Gomez Ortega et al., 2021). This distinguishes them from methods such as data linkage. For example, when accessing Twitter data via linkage, participants only provide their user handles in a survey, allowing researchers to access public profiles and additional information (Silber et al., 2022). In contrast, data donation enables participants to share Twitter DDPs, which contain more detailed and non-public information (Boeschoten et al., 2022; Silber et al., 2022; Strier et al., 2020). Some studies also allow participants to review and delete data before donating, further increasing active involvement (e.g., Breuer et al., 2022; Zannettou et al., 2023).

Additionally, data donations are 2) *unpaid*, as participants are not directly compensated for their data (Susha et al., 2019, p. 118). While incentives may be offered to compensate time and participation, they do not constitute payment for the data itself and thus preserve the notion of donation for the common good (Bietz et al., 2019, p. 8).

Next, data is 3) *not generated for research* (Bietz et al., 2019, p. 1). This study focuses on trace data, meaning historical records of participants' digital behavior (Pak et al., 2022, p. 135). Examples include data from fitness trackers (Pilgrim & Bohnet-Joschko, 2022), smartphone usage statistics (Baumgartner et al., 2023;

Ohme et al., 2021), social media data (e.g., Boeschoten et al., 2022; Pak et al., 2022) or Google search histories (Blassnig et al., 2023). Trace data is generated by systems and reflects user behavior (Howison et al., 2011), offering a more objective account than self-reports (Ohme et al., 2021). To qualify as a valid donation, data must be captured non-reactively, meaning it is not self-reported but provided as system-exported files (e.g., DDPs) or other materials containing recorded behavior, such as screenshots.

The next facet concerns the 4) *intangible and reproducible nature* of data donations, reflected in “data multiplicity” (Prainsack, 2019, p. 13). Unlike physical donations, data can be shared multiple times without loss to donors. However, donating data may feel less rewarding, as other forms of donation, such as blood donation, can evoke a ‘warm glow’ (e.g., Skatova & Goulding, 2019, p. 3).

Another aspect is that data is donated 5) *retrospectively*, meaning consent is given ex-post (Stier et al., 2020, p. 506), for data that already exists. The data were generated in the past, logged by platforms or services, and donated later. This distinguishes data donations from other user-centric trace data collection methods, such as tracking, where participants provide consent ex-ante and install tools that continuously collect data (Gil-López et al., 2023; Stier et al., 2020, p. 506). Retrospective donation can help mitigate reactivity during data collection (e.g., Baumgartner et al., 2023, p. 5).

Further, participants donate their data 6) *voluntarily and with informed consent* (e.g., Bietz et al., 2019, p. 1). They are informed about what data is requested, the procedures involved, and how the data will be handled, analyzed, and shared.

Finally, data donations are 7) *intended to serve the common good* and support societal well-being, as they are typically motivated by research rather than financial interests (Bietz et al., 2019; Hillebrand et al., 2023). However, the societal value of donated data depends on the type of research and, in particular, on researchers’ ethical and responsible use of the data.

Summarizing, this study defines data donations as: 1) an *active* and 2) *unpaid* transfer of personal 3) *trace data not generated* for research to researchers. The data is 4) *intangible and reproducible*, and collected 5) *retrospectively*. Consent is 6) *voluntary and informed*, and 7) the study is intended to *serve the common good*.

To illustrate our definition more clearly, we provide examples of two data types and research designs that do not align with it and are therefore excluded from this study, namely genomic data and posthumous data sharing. First, our definition requires active data transfer by participants. In many cases, genomic data is already held by researchers or medical professionals, and participants only consent to its use, resembling data linkage rather than active donation. Second, our focus is on trace data reflecting past behavior. Although genomic data is highly personal, it does not capture behavior and is therefore not relevant to this review. While it can originate outside research contexts, such as medical diagnostics or ancestry services, it is often collected specifically for research, whereas our focus is on data generated independently of research. Third, unlike the fourth aspect of our definition, genomic data collection often involves a physical component, such as providing tissue samples (e.g., Mezinska et al., 2020, p. 3). This entails physical

loss and potential health risks, distinguishing it from the intangible nature of data donations as defined here.

A second excluded example is posthumous data provision. In this case, consent is not given retrospectively but earlier in life, when not all data donated at death may yet exist. This violates the fifth part of our definition.

2.2 Research on participation in data donation studies

Participation bias in data donation studies is difficult to assess due to varying definitions and uses of the term across disciplines. While common in communication research (e.g., Boeschoten et al., 2022; van Driel et al., 2022) and health research (e.g., Bietz et al., 2019), it also appears in fields such as human-computer interaction (e.g., Gomez Ortega et al., 2021), often with limited cross-disciplinary awareness.

Literature on health-related data donations is most prominent in research (Susha et al., 2019) and reports high rates of willingness to participate, ranging from 65% (Seltzer et al., 2019, p. 2) to 84% (Buhr et al., 2022, p. 9). In contrast, in communication research, the likelihood of donation tends to fall below the midpoint of a scale (e.g., Pfiffner & Friemel, 2023, p. 15) and donation rates are often quite low, around 20% (Hase & Haim, 2024; Keusch, Pankowska, et al., 2024; Strycharz et al., 2024). Additionally, different types of data might be prone to different perceptions and biases (e.g., Pak et al., 2022; Pfiffner & Friemel, 2023), as health data is perceived as more sensitive than social media data (Silber et al., 2022, p. 401).

Altogether, findings on (non-)participation in data donation studies are often not integrated. Thus, this review seeks to provide a comprehensive overview of research, mapping factors that may influence (non-)participation. By including multiple disciplines, it offers a broad perspective and insights beyond social media donations. Therefore, we ask:

RQ1: What characterizes studies on willingness to donate (e.g., research fields, methodology, hypothetical vs. actual donation)?

2.3 What factors might influence data donations

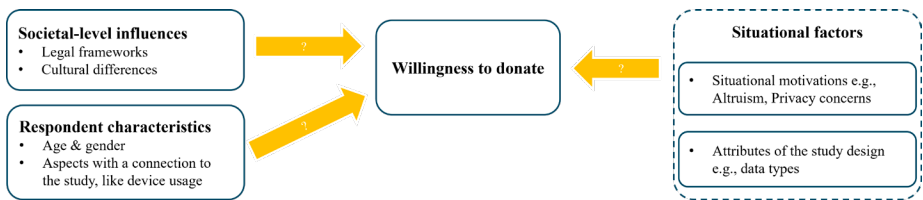
To examine participation in data donation, we draw on Groves et al.'s (1992) framework of survey non-response. It distinguishes three factors: (1) characteristics of the sample person, (2) societal-level factors, and (3) attributes of the survey design. As data donation studies typically lack face-to-face interaction, we exclude interviewer-related aspects such as interviewer attributes and respondent–interviewer interaction.

First, we adopt the category ‘characteristics of the sample person’, which we term *respondent characteristics*. These are stable demographic variables, such as age or gender, that are independent of the data donation context. Second, we examine *situational influences*, including ‘attributes of the survey design’ as adopted from Groves et al. (1992) as one sub-dimension. This covers factors such as the

wording of the data request, which can affect willingness to donate (e.g., Silber et al., 2024). As a second sub-dimension, we add *situational motivations*, such as privacy concerns. Unlike stable demographic variables, these can vary depending on the situation and perceived risks or circumstances

Finally, we consider ‘*societal-level factors*’, referring to broader influences such as legal frameworks (e.g., GDPR), national culture, or major events like health crises. These factors shape general attitudes toward research and can influence data-sharing behavior. Together, these categories structure our analysis of participation (see Fig. 1) and are discussed in the following sections.

Figure 1. Categories of analysis



2.3.1 Respondent characteristics influencing non-participation

First, we lay a focus on *respondent characteristics* that influence participation, also referred to as dispositional variables (Strycharz et al., 2024). Survey research shows that such characteristics, including sociodemographic traits like gender (Keusch, 2015, p. 189) and personal interest in the topic (p. 186), can affect participation in research.

Similarly, research on data donation has observed differences between donors and non-donors, for example, regarding political interest (Hase & Haim, 2024, p. 17), age and digital and algorithmic efficacy (Strycharz et al., 2024, p. 15), as well as general trust in research (Keusch, Pankowska, et al., 2024, p. 8) Therefore, we ask:

RQ2: Which respondent characteristics (e.g., sociodemographic characteristics, device usage) are discussed as potential determinants of willingness to donate?

2.3.2 Situational factors influencing non-participation

Situational factors influencing non-participation can be divided into two subcategories: *Attributes of the survey design* and *situational motivations*. According to Groves et al. (1992, p. 477–478), ‘*attributes of the survey design*’ refer to study characteristics and design choices that affect participation decisions. This factor is also relevant for data donation studies, as elements such as the type of data requested can influence willingness to donate (Kmetty et al., 2024; Silber et al., 2024).

We also assume that *situational motivations* are relevant. In data donation studies, participants share personal data, so privacy concerns may influence partici-

pation (Boeschoten et al., 2022; Ohme et al., 2021; Pfiffner & Friemel, 2023). The Privacy Calculus framework helps explain this behavior, suggesting that individuals weigh potential benefits and risks when deciding whether to disclose information, though decisions may not always be rational or consistent with general attitudes (Aquisti, 2004; Dinev et al., 2015).

This suggests that privacy concerns are not stable traits but context-dependent states. Decisions are made within specific situations and thus reflect situational motivations. Similarly, tendencies such as altruism, often considered a stable trait, may also depend on context. Skatova and Goulding (2019) show that motivations for data donation extend beyond general altruism and are more context-dependent. Accordingly, this study emphasizes situational factors influencing participation in data donation.

RQ3: What situational factors (i.e., situational motivations and attributes of the study) are discussed as influencing willingness to donate?

2.3.3 Societal-level influences on participation

Finally, Groves et al. (1992) identify ‘societal-level factors’ as influences on survey participation, referring to the broader societal context of a request (p. 477). In data donation, cultural differences may shape openness to data sharing across populations. Legal frameworks such as the GDPR can influence both the feasibility of such research and public awareness of data access. Broader events, such as health crises, may also affect attitudes toward research and willingness to donate data. Consequently, the following research question is posed:

RQ4: What societal-level factors (e.g., cultural aspects, jurisdiction) are discussed as influencing willingness to donate?

2.3.4 Which strategies might enhance participation?

Research on data donations has examined not only predictors of non-participation (Pak et al., 2022; Pfiffner & Friemel, 2023; Silber et al., 2022; Strycharz et al., 2024; Welbers et al., 2024) but also strategies to increase participation. In particular, ex-ante strategies targeting survey design are a key focus (de León et al., 2025; Hase & Haim, 2024). One approach is monetary incentives, which can be effective for certain data types (Kmetty et al., 2024; Silber et al., 2022) though they may introduce bias (Zannettou et al., 2023, p. 8).

Effective strategies must address the specific benefits and concerns of data donation, and some studies have begun to explore these (de León et al., 2025; Hase & Haim, 2024; Keusch, Pankowska, et al., 2024). For example, researchers may offer personalized feedback or insights about participants’ data (Gomez Ortega et al., 2021, p. 498). Addressing privacy concerns is also essential, for instance, through strong data protection measures (e.g., Boeschoten et al., 2022, p. 408).

Nudges have also been proposed to increase survey participation. They are defined as “any aspect of the choice architecture that alters people’s behavior in a

predictable way without restricting options or significantly changing economic incentives” (Thaler & Sunstein, 2008, p. 6). Like data donations, nudges aim to promote behavior that benefits individuals or society.

One example is framing, used in survey design to increase willingness to donate data. This can involve emphasizing the societal value of data donation (Hase & Haim, 2024) or highlighting the potential loss of valuable data if individuals do not donate (Keusch, Pankowska, et al., 2024). Although framing is not always effective, nudges remain a promising strategy. This leads to our final research question:

RQ5: What strategies (i.e., general or relating to study design) are proposed in the literature to increase participation willingness?

3. Method

To address the research questions, we conducted a scoping review, a type of literature review used to map and organize diverse research fields (Arksey & O'Malley, 2005; Levac et al., 2010). Scoping reviews capture the breadth of research and identify overarching themes and trends (Arksey & O'Malley, 2005, p. 21). They also enable the integration of fragmented findings and the identification of factors influencing a concept (Peters et al., 2020, p. 2122). This approach is therefore well suited to examining factors influencing willingness to donate across disciplines. A scoping review is more appropriate than a systematic review for this study, as systematic reviews require narrowly defined questions on constructs, effect directions, or specific designs (Munn et al., 2018; Peters et al., 2020). This would be difficult given the diversity of research on willingness to donate and the exploratory focus of this study.

A research protocol (Booth et al., 2022; Peters et al., 2020; see online supplement 1 in OSF) was developed in advance according to PRISMA-P guidelines (Moher et al., 2015). The following section reports required elements and deviations from the protocol in line with PRISMA-ScR guidelines (Tricco et al., 2018).

3.1 Outcomes

The primary outcomes aim to address the research questions by identifying factors examined in the literature, such as RQ2) respondent characteristics, RQ3) situational influences, RQ4) societal-level influences on willingness to donate and RQ5) strategies to enhance participation. To address RQ1, we also examine the scope of the literature across disciplines, trends over time, types of data donated, and described donation processes.

3.2 Deviations from the protocol

We were unable to fully adhere to the research protocol, as outcomes such as actual donation and compliance rates were omitted. This was due to the inclusion of many hypothetical studies, where such measures were not applicable. Additio-

nally, willingness to donate was measured inconsistently across studies, for example, through categorical responses (“yes”, “no”, “don’t know”; e.g., Buhr et al., 2022) or continuous scales from 0 to 100% (Hillebrand et al., 2023), limiting comparability. No measures of strategy effectiveness were reported.

Finally, the research question on situational influences was divided into two, placing greater emphasis on societal-level factors as a distinct category.

3.3 Eligibility criteria

To be included, studies had to be published between January 1, 2018, and April 17, 2023, when data collection occurred. The cutoff reflects the introduction of the GDPR, which likely facilitated data donation research. In the same year, iOS and Android introduced phone usage statistics (Android.com, 2018, September 25; Apple Inc., 2018, September 25), enabling access to behavioral data that can be donated, for example, via screenshots (Ohme et al., 2021).

Studies were selected based on inclusion criteria aligned with our definition: 1) an active and 2) unpaid transfer of personal 3) trace data not generated for research to researchers. The data must be 4) intangible and reproducible, 5) collected retrospectively, with 6) voluntary and informed consent, and 7) an aim to serve the common good.

To focus on literature relevant to the research questions, we excluded studies on 1) feasibility, 2) ethics, 3) legal reviews, 4) technical frameworks, and 5) applications of data donation without a focus on participation or bias. To comprehensively understand factors influencing willingness to donate, both empirical and theoretical work were considered. Grey literature was not systematically searched, but was included if identified.

3.4 Search strategy

To identify primary literature, we conducted database and Google Scholar searches using the same terms. The databases included 1) Web of Science, 2) ScienceDirect, 3) EBSCOhost (Business Source Complete, Academic Search Complete, Communication & Mass Media Complete, APA PsycInfo, APA PsycArticles, and Information Science & Technology), and 4) ACM Digital Library. Search results were downloaded as BibTeX files for screening. For Google Scholar, only the first five result pages were considered due to the high number of results (Haim et al., 2018).

To identify secondary literature, we consulted an expert (e.g., Booth et al., 2022), who suggested relevant studies after being briefed on the review’s goals. We also conducted a backward citation search (Peters et al., 2020) by screening references of included studies and reviewing abstracts of unfamiliar but potentially relevant sources.

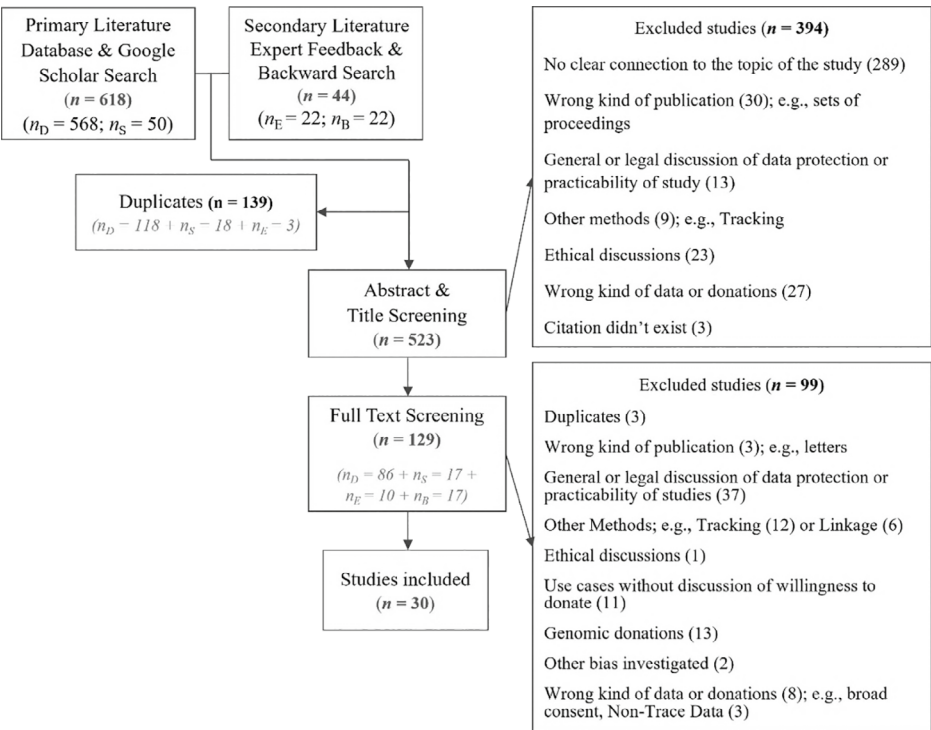
Search terms were developed in English and German and refined based on expert feedback: “Data don*” OR (donat* AND “Digital Trace”) OR (“usercentric” AND donat*) OR (“user-centric” AND donat*) OR (“user centric” AND donat*) OR Datenspend* OR “Daten spend*”. Where truncation was not possible, database-specific filters were applied.

3.5 Selection process

The first stage of the selection process (Fig. 2), abstract screening, was conducted using Rayyan.ai (Kellermeyer et al., 2018). BibTeX citations from database searches were uploaded, and papers were manually labeled as included for full-text screening or excluded based on their abstracts. Missing abstracts were retrieved and screened manually. Additional features of Rayyan.ai¹, such as word clouds, were not used. Google Scholar results and secondary literature were documented in a Microsoft Word file, where eligibility decisions were also recorded.

In the second stage, full texts were screened in Citavi², and eligible studies underwent backward citation screening. After the selection process (Fig. 2), 30 articles were included (see online Appendix 2). Exclusions labeled “Citation didn’t exist” refer to three database entries for which no corresponding publications could be identified.

Figure 2. Selection process



Notes. n_D = Database; n_S = Google Scholar; n_E = Expert feedback; n_B = Backward search

1 <https://www.rayyan.ai/>
2 <https://www.citavi.com/de>

3.6 Data charting

Data charting included both empirical findings and theoretically discussed factors and strategies reported in the papers. Categories were predefined (online supplement 1 in OSF) and revised after charting the first five articles.

We excluded categories such as goals or research questions, population, sample size, key findings, strategy effectiveness, donated data, initial response rate, and compliance rate, as these outcomes were often unavailable due to the hypothetical nature of many studies. We added the categories “type of publication” and “applied strategies or incentives to enhance participation,” and refined others to better capture the construct of interest (online supplement 3 in OSF).

Study fields were coded by the first author, primarily based on the journals. When unclear, the paper’s main topic and author affiliations were considered, and a second field was assigned when necessary.

3.7 Data analysis

The analysis strategy was not specified in the research protocol due to the unpredictable diversity of the data. We conducted an inductive thematic analysis (Braun & Clarke, 2006), a method used in scoping reviews (e.g., Severin-Nielsen, 2023) to identify themes, here understood as influencing factors. Relevant text passages were abstracted, summarized, and consolidated into overarching factors (Tab. 1). These were included as subcategories to describe the identified factors (online supplement 4 in OSF).

Table 1. Example of data analysis

	Abstraction	Factor	Subcategories
“Patients also expressed concerns about potential data and privacy breaches” (Seltzer et al., 2019, p. 3)	Privacy concerns	Privacy concerns	Security concerns, negative experiences
“Willingness toward sharing PHD is related with privacy and security considerations.” (Karampela et al., 2019, p. 9)	Privacy & security concerns		
“Privacy-Related Experiences: - Was the victim of fraud and/or identity theft” (Hendricks-Sturup et al., 2022, p. 5)	Privacy-related negative experiences		

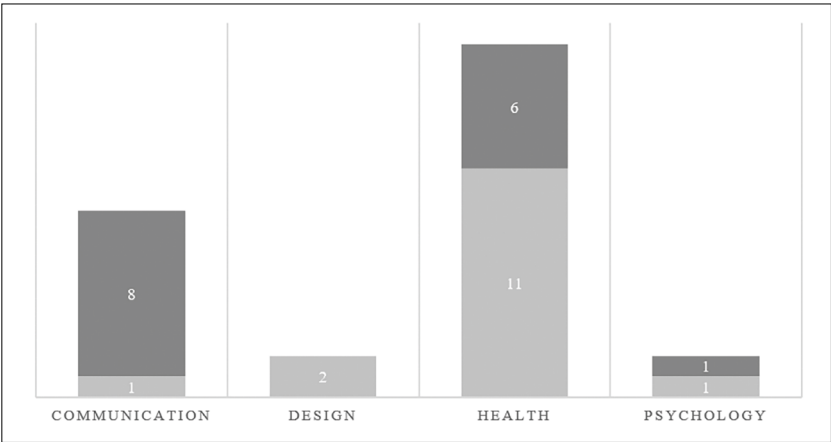
4. Results

4.1 Descriptive overview of studies' characteristics

Of the 30 publications, 26 are journal articles, three are conference papers (Diethei et al., 2021; Gomez Ortega et al., 2021; Maus et al., 2020), and one is an essay (Bietz et al., 2019). Online appendix 5 details study categorizations.

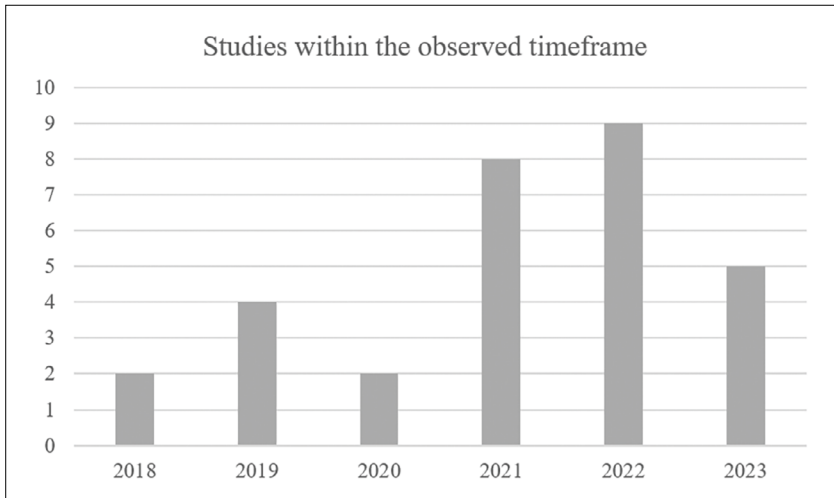
Regarding RQ1, which examines disciplinary background, methodology, and focus on hypothetical versus actual donations, health research dominates the sample ($n = 17$), followed by communication ($n = 9$), which often has a methodological focus.

Figure 3. Disciplines of included studies



Note. The upper section of the bar in darker color indicates studies with a methodological focus.

Examining the publication years (Fig. 4), an upward trend up to 2023 can be shown. The decrease seen in 2023 is likely due to data collection ending in April.

Figure 4. Publications over time

Note. The timeframe investigated for this study lays within January 2018 to April 2023

Regarding methodological approaches (Tab. 2), surveys are most common ($n = 12$), followed by actual data donation studies ($n = 9$), interviews ($n = 3$), theoretical discussions ($n = 3$), experiments ($n = 2$), and focus groups ($n = 1$). Table 2 provides a more detailed breakdown, with the “Total” row indicating how many studies primarily used each method.

Table 2. Detailed classification of methodological approaches included studies used

	Data donations	Experiment	Qualitative interview	Survey	Theoretical work	Focus groups
+ Theoretical discussion	1					
+ Diary	1					
+ Survey	5					
–	2					
–		2				
+ Content analysis			1			
+ Participant observation			1			
–			1			
–				12		
+ Discussion group					1	
–					2	
–						1
Total	9	2	3	12	3	1

Notes. “+” indicates the method in the column was combined with the one in the row; “–” means no additional method was used. E.g., “5” in the *Data Donation* column and “+Survey” row means five studies primarily used data donation and also included a survey. The “2” in the *Data Donation* column and “–” row shows two studies only using data donations.

Further analysis shows that most studies ($n = 18$) examine hypothetical willingness to donate data. This is followed by actual donations ($n = 7$) and purely theoretical perspectives ($n = 3$) and studies investigating hypothetical scenarios along with an actual donation ($n = 2$).

Investigating the kinds of data the studies (hypothetically) requested (Tab. 3), we can see that health research dominates once more.

Table 3. Data types included studies investigated

Investigated data types	<i>N</i>
Health data	18
Social media	9
Finance data	4
Google searches	4
Other (e.g., ‘machine data’, voting history)	3
Smartphone usage (e.g., battery, screen-time)	3
Mobility (e.g., ride sharing, GPS data)	2
Musik streaming	2
Communication (e.g., mail records, text messages)	2

Note. One study can investigate more than one data type.

Concerning the donation processes (Tab. 4), the precision of questions and descriptions varies. Most studies ($n = 14$) do not explicitly state which donation process they assume, often just asking for willingness to donate a certain data type (e.g., Pilgrim & Bohnet-Joschko, 2022) or concerns when donating data (Hendricks-Sturup & Lu, 2023).

Table 4. Donation methods investigated

Investigated donation method	<i>N</i>
Unspecified	14
Data upload through websites or surveys	6
Specialized tool (DDPs)	5
Through apps	5

Note. One study can investigate more than one donation process.

4.2 Respondent characteristics

Table 5 shows how many studies discussed respondent characteristics as potentially influential on willingness to donate, addressing the second research question, and showcases exemplary subcategories. Online appendix 5 shows assignments, which studies count into which category.

Table 5. Respondent characteristics influencing willingness to donate

	Emp. Theor.		Exemplary subcategories
Sociodemographic	13	2	Age, gender, education, ethnicity, migration background, income
Technical skills	3	6	Media & technology literacy, technological affinity
Device usage	6	2	Frequency of App use, platform usage
Characteristics related to the study	5	2	Politics: Political interest & political efficacy Other: Health status, news consumption
Other characteristics	5	0	Trust, Big Five personality traits, other donation behavior
Research compliance	2	2	Participation in experiments, attitudes to surveys
Device ownership	4	0	Type of device, operating system

Notes. One study can investigate more than one factor. Column ‘emp.’ Summarizes the number of studies empirically investigating the concept, ‘theor.’ theoretical discussions.

First, sociodemographic factors such as gender, age, and income are frequently discussed as influencing willingness to donate ($n = 15$). Hypothetical studies often report negative effects of age (Hillebrand et al., 2023; Karampela et al., 2019; Kaspar, 2020; Pilgrim & Bohnet-Joschko, 2022). When gender effects are significant, men tend to be more willing to donate data in both actual (e.g., Silber et al., 2022) and hypothetical contexts (Karampela et al., 2019; Pilgrim & Bohnet-Joschko, 2022).

Technical skills ($n = 9$) and device usage ($n = 8$) are also commonly identified as relevant respondent characteristics.

Study-related characteristics ($n = 7$) further influence participation. For example, individuals with a disease may be more willing to contribute to related research (Karampela et al., 2019). Physical activity can also matter, as individuals may hesitate to share fitness data if they perceive themselves as insufficiently active or not fitting certain health ideals (Toepoel et al., 2021, p. 4).

Finally, other personal characteristics ($n = 5$), such as Big Five traits, general research compliance ($n = 4$), and device ownership ($n = 4$), are also linked to willingness to donate.

4.3 Situational factors

Table 6 provides an overview of situational motivations that potentially influence willingness to donate. This is complemented by Table 7, which summarizes the findings on the attributes of the study design. Together, they answer the third research question.

Table 6. Situational motivations influencing willingness to donate

	Emp. Theor.		Exemplary subcategories
Privacy concerns	18	4	Concerns on data storage, negative privacy related experiences
Egoism	10	3	Positive: Direct or future benefit Negative: Direct loss
Altruism	10	2	General: Desire to help others and contribute to social welfare Collectivism: Benefit for direct community Impure altruism: Indirect reciprocity Reluctant altruism: Perceived need for donation

Notes. One study can investigate more than one factor. Column ‘emp.’ Summarizes the number of studies empirically investigating the concept, ‘theor.’ theoretical discussions

Most studies identify privacy concerns ($n = 22$) as a key factor influencing willingness to donate. These include concerns about data use and prior experiences with privacy breaches (e.g., fraud). Most studies find a negative effect on actual donation (e.g., Ohme et al., 2021; Pak et al., 2022) or willingness to donate (e.g., Hillebrand et al., 2023; Karampela et al., 2019). Some report effects only for specific data types or inconclusive results (Hendricks-Sturup et al., 2022), likely due to varying operationalizations.

Egoism ($n = 13$) is also frequently discussed. Participation may increase when personal benefits are expected but decrease when participants anticipate risks or losses, such as concerns about consequences of sharing online searches.

Altruism ($n = 12$) is another key factor with several forms. These include pure altruism, the desire to contribute to the common good (Skatova & Goulding, 2019), and collectivism, the intention to support one’s community (Diethei et al., 2021, p. 7). Impure altruism reflects self-interested motives, such as expectations of indirect reciprocity, where individuals believe others will also donate, enabling mutually beneficial research (Hillebrand et al., 2023; Skatova & Goulding, 2019). This aligns with findings that individuals are more willing to share data when they expect others to do so (Pilgrim & Bohnet-Joschko, 2022, p. 9). Finally, reluctant altruism occurs when individuals donate because they perceive a need and believe others will not contribute (Skatova & Goulding, 2019, p. 3).

Table 7. Attributes of the study design influencing willingness to donate

	Emp.	Theor.	Exemplary subcategories
Data types	8	4	Sensitivity of data, data granularity, perceived relevance of data
Recipient of the data	9	2	Trust in recipient of the data, attitudes towards recipient of the data, type of institution receiving data
Understanding & impact	9	1	Understanding: Perceived value of study for society or research Impact: Perceived impact of one's contribution, uncertainty about value of contribution
Conditions of sharing	8	1	Recipient of the data: Commercial vs. academic Donation process: Type of donation process, trust in process
Purpose of the study	5	2	Attitudes towards the purpose of the study
Control in the process	4	1	Autonomy and sovereignty, perceived behavioral control
Transparency	3	1	Accountability for data and research
Technical problems	1	3	
Presentation & explanation	1	2	Question positioning and posing, evaluation of survey
Effort of participation	0	3	Forgetting to participate
Informed consent	2	1	Understanding the consent, type of consent

Notes. One study can investigate more than one factor. Column ‘emp.’ Summarizes the number of studies empirically investigating the concept, ‘theor.’ theoretical discussions

A key study design factor is the type and sensitivity of requested data ($n = 12$). Health data is often perceived as more sensitive than social media data, reducing willingness to donate (Silber et al., 2022). However, perceived relevance also matters, as the research value of health data may be more apparent to participants. Trust in the data recipient ($n = 11$) is another important factor.

Understanding and perceived impact ($n = 10$) also influence participation. Participants may be more willing to donate if they understand the study’s purpose and believe their contribution is meaningful, highlighting the importance of clear communication.

Willingness further depends on sharing conditions ($n = 9$), such as whether researchers are from academic or industry contexts and how trustworthy the dona-

tion process is. The study's purpose ($n = 7$) also matters, as positive attitudes toward it can increase participation.

Control over the donation process ($n = 5$) is another factor. Allowing participants to review or delete data can enhance perceived control and encourage participation.

Less frequently discussed factors include transparency ($n = 4$), technical issues ($n = 4$), and clarity of presentation and explanation ($n = 3$). Effort required to participate ($n = 3$) may also reduce motivation or lead to non-participation. Finally, informed consent ($n = 3$), including clarity and potential time constraints, can influence willingness to donate.

4.4 Societal-level influences

Societal-level influences played a minor role in literature. However, those that were discussed are summarized in Table 8 in response to the fourth research question.

Table 8. Societal-level influences on willingness to donate

	Emp. Theor.		Exemplary subcategories
Contextual sensitivity	3	3	Relevance of time of (health-) crisis, concrete concerns for the community
Country & culture	2	1	East vs. West Germany, rural vs. urban, degree of digitalization of the country

Notes. One study can investigate more than one societal-level influence. Column 'emp.' Summarizes the number of studies empirically investigating the concept, 'theor.' theoretical discussions

We identified contextual sensitivity ($n = 6$) as a factor influencing willingness to donate. In particular, crises, such as pandemics, can increase participation, as heightened health concerns may shape behavior.

Country and culture ($n = 3$) also emerged as relevant influences. Factors such as levels of digitalization and whether individuals live in rural or urban areas may affect willingness to donate data.

4.5 Strategies

Finally, answering the fifth research question, we want to shed light on strategies that might augment willingness to donate, once more divided into general and those specifically relating to study design (Tab. 9 & 10).

Table 9. General strategies to augment willingness to donate

		<i>n</i>	Exemplary subcategories
Benefits	Offering benefits	5	Insight into data, visualizations
	Monetary incentives	7	Sweepstakes, direct payments
	Community	2	Establishing community for exchange or comparison
Communication	Address altruism	5	Emphasizing the societal relevance of study
	Emphasize the relevance	4	Providing information on value of contribution, emphasizing the need for donation, applicability of study results for the greater good
	Express appreciation	2	Involving participants in the research process, acknowledging participants contribution
Nudges	Reminder	3	Sending reminders
	Framing	2	Framing contribution of data as a donation, pro-social framing
	Individualizing messages	1	Addressing participants motivations
	Timing	1	Posing request after participants already shared some data, event-based requests
Manipulation	Personal recruitment	1	Utilizing the higher compliance typically observed in in-person interactions
	Hidden request	1	Intransparent requests leading to inadvertently or unknowingly agreement to participate

Note. One study can investigate more than one strategy.

To increase willingness to donate, studies may offer benefits ($n = 4$), appealing to egoistic motivations. These include personalized products (Gómez Ortega et al., 2021, p. 498) and monetary incentives ($n = 7$), though optimal incentive levels remain unclear (e.g., Silber et al., 2022). Establishing communities ($n = 2$) for sharing or comparing donated data has also been suggested.

Effective communication is essential. Emphasizing study relevance ($n = 4$) and appealing to altruism ($n = 5$) can encourage participation, while expressing appreciation ($n = 2$) may target egoistic motives.

Nudges offer additional strategies, such as reminders ($n = 3$) to reduce forgetfulness and framing ($n = 2$) to activate altruistic motivations. Other approaches include tailoring messages to individual motivations ($n = 1$) and timing requests strategically ($n = 1$), for example, when participants are already sharing data.

Some studies also discuss manipulative techniques, such as personal recruitment ($n = 1$) or hidden requests ($n = 1$), though these raise ethical concerns (Pilgrim & Bohnet-Joschko, 2022, p. 8).

Finally, strategies related to study design (Tab. 10) are closely linked to the identified design attributes influencing willingness to donate (Tab. 7).

Table 10. Strategies relating to study design attributes to augment willingness to donate

	<i>n</i>	Exemplary subcategories
Cultivating trust	14	Collaborations with trustworthy partners, no commercial use of data
Privacy protection & emphasis	10	Local processing, data minimizing, anonymizing, respecting legal guidelines
Reducing burden & effort	9	Execution: Offering instructions and technological support Preparations: Embedded requests, simplifying processes & interfaces
Autonomy & control	8	Data types: Offering participants the possibility to filter and delete parts of their data Other: Usage of familiar devices, possibility to drop out
Informed consent	7	Intelligibility of consent: Participants should understand consent forms Appropriateness of consent: Transparent and potentially selective consent
Data types	5	Asking for less granular and sensitive data

Notes. One study can investigate more than one strategy. Column ‘emp.’ Summarizes the number of studies empirically investigating the concept, ‘theor.’ theoretical discussions

First, cultivating trust ($n = 14$) is widely recommended, for example, through collaboration with trusted partners (Diethei et al., 2021, p. 11) or involvement of ethics boards (Boeschoten et al., 2022, p. 409). Trust improves perceptions of data recipients and transparency and can mitigate privacy concerns (Juga et al., 2021).

Given the prevalence of privacy concerns, strong privacy protection and clear communication of safeguards are essential ($n = 10$). Aligning measures with participants’ expectations can increase trust and willingness to donate.

Reducing effort and burden ($n = 9$) is also important to prevent dropouts. This includes simplifying procedures and offering technical support to address limited skills.

Increasing autonomy and control ($n = 8$) is another key strategy. Allowing participants to review, select, or delete data can enhance control, awareness, and digital literacy (Buhr et al., 2022, p. 10; Gómez Ortega et al., 2021, p. 497). However, greater perceived control may improve attitudes toward sharing while reducing willingness to share (Juga et al., 2021, p. 53). Clear and appropriate informed consent is also recommended ($n = 7$).

Finally, requesting appropriate data types ($n = 5$), particularly less sensitive data, can increase willingness to donate. While research needs may limit this, data

minimization should be considered (Hakobyan et al., 2025; Ohme & Araujo, 2022, p. 1).

5. Discussion

The aim of this scoping review was to examine research on participation in data donation studies and factors influencing willingness to donate.

First, we analyzed what characterizes the literature on willingness to donate. To date, the number of studies is relatively small, with more than half of the studies under investigation originating from a health context. This is important, as concerns about data donation may vary across fields. In health contexts, data is often perceived as particularly sensitive, but its potential value and benefits for others (Kulzer et al., 2022) may also be more salient.

Most studies examine hypothetical willingness to donate and use a variety of measurement approaches. This reliance on hypothetical designs is a limitation, as such studies may overestimate actual participation rates. However, since this study was conducted, more research has begun to include actual data donations (e.g., Keusch, Pankowska, et al., 2024; Strycharz et al., 2024). While surveys provide insight into attitudes, stated intentions do not always translate into behavior. Research on the privacy paradox (e.g., Norberg & Horne, 2007) highlights this gap. For example, individuals often report concerns about sharing data with companies but still do so for minor benefits. In the context of data donation, the opposite pattern may occur: Individuals may express support for donating data for research but refrain from doing so, possibly due to the absence of immediate personal benefits.

When measuring hypothetical willingness to donate, wording is crucial. A lack of specificity about the donation process can influence responses. Some studies ask about willingness without clarifying how or under what conditions data would be donated, which is problematic given that study attributes affect decisions. For example, the method of data collection, such as via a platform or questionnaire, can lead to different levels of willingness (Toepoel et al., 2021, p. 8).

Second, we examined respondent characteristics. Studies most often focus on sociodemographic factors, with men generally more likely to donate, as well as technical skills and device usage. Study-related characteristics are also relevant. For instance, individuals may choose not to donate fitness data due to low activity levels (Toepoel et al., 2021, p. 4), which may introduce social desirability bias in participation decisions rather than in the data itself.

Regarding the third research question on situational factors, privacy concerns, egoism, and altruism emerge as key motivations. However, general altruism alone has limited explanatory power, as other mechanisms and forms of prosocial behavior also shape participation (Skatova & Goulding, 2019, p. 12).

Altruism can take different forms, including impure, reluctant, and collectivist altruism. The latter is particularly relevant, as individuals may struggle to see how their contribution benefits broader society. Framing donations as supporting an immediate community can make participation more tangible. For example, indi-

viduals who have lost someone to suicide may better recognize the value of their contribution when it supports prevention research (Sleigh, 2018, p. 4).

Regarding study design attributes, the literature highlights data types, the data recipient, and participants' understanding of the study and its impact. These can be interpreted through the Theory of Contextual Integrity (Nissenbaum, 2010), which conceptualizes privacy as appropriate information flow based on context-specific norms. These norms involve actors (e.g., recipients), attributes (e.g., data types), and transmission principles (e.g., conditions of sharing). Willingness to donate may therefore depend on whether study attributes align with participants' expectations of appropriate data flows, which can also foster trust.

Our findings suggest several practical recommendations for designing data donation studies. Researchers should follow established best practices to ensure ethical and methodological standards (Carrière et al., 2024; Ohme & Araujo, 2022). At the same time, they must carefully select and prepare requested data while considering structural constraints, particularly those related to DDPs, whose content and format are controlled by platforms (Hase et al., 2024).

To address technical difficulties and reduce effort, lowering barriers to participation should be a priority. This includes clear instructions and support services, such as help desks, to guide participants.

Aligning data requests with participant preferences, for example, by avoiding highly sensitive data, can build trust and support data minimization. Whitelists can help achieve this (Hakobyan et al., 2025; Ohme & Araujo, 2022). Allowing participants to remove parts of their data can further increase perceived control and autonomy.

If the study design cannot be adapted due to specific data requirements, communication becomes even more important. Transparency about data handling can help address privacy concerns, although it may also increase their salience (Marreiros et al., 2017, p. 10). Communicating the non-commercial nature of the research and adhering to ethical standards can strengthen trust. Emphasizing the societal relevance of the study and the impact of individual contributions can further increase willingness to participate. Offering meaningful incentives, such as personal data insights, may also enhance motivation.

Furthermore, nudges may be an effective strategy to increase willingness to donate data, for example, through reminders or careful timing of requests. In contrast, manipulative strategies such as hidden requests can be considered forms of "sludge," meaning illegitimate interventions that provide no benefit to individuals or society (Thaler, 2018, p. 431). Such practices should be avoided, as they may undermine autonomy and comfort, erode trust in research, and reduce willingness to participate in the future. Study design strategies should therefore prioritize transparency, non-manipulation, and responsible use of behavioral interventions.

Altogether, strategies to increase willingness to donate require further investigation. While they cannot address all biases inherent in data donations, such as sampling bias from hard-to-reach populations, they can improve participation among accessible groups. If the goal is to increase response rates within these groups, the strategies and factors outlined here should be considered. However,

there is no universal solution, and not all ex-ante strategies are effective in every context (Hase & Haim, 2024; Keusch, Pankowska, et al., 2024). Complementary approaches, such as retrospective statistical correction, may therefore be valuable (Pak et al., 2022).

It remains important to identify which strategies are effective in specific contexts to encourage donations. This is particularly relevant in communication research, where data donations enable the study of digital behavior and key societal issues, such as algorithmic personalization and online participation, that are difficult to examine accurately with other methods.

Future research requires a more consistent understanding of variables such as technical skills and willingness to donate. However, these measures must remain context-sensitive. For example, the skills needed for screenshot donations differ from those required for handling DDPs, so measurement tools should reflect these specific demands. Developing comparable measures for each task would improve consistency across studies.

Furthermore, privacy concerns also vary by data type and sensitivity. Concerns about private chat data in DDPs may center on potential data leaks, whereas concerns about social media data may relate to repercussions for expressed opinions.

This study has several limitations. First, data donation is a relatively new research method, and the field is evolving rapidly. Since this study was conducted, additional literature has emerged, particularly in communication research (e.g., de León et al., 2025; Haim & Hase, 2024; Keusch, Pankowska, et al., 2024; Silber et al., 2024; Strycharz et al., 2024). The review should therefore be seen as a snapshot of the field at a specific point in time. As research continues to grow, updating this work or conducting full reviews may prove valuable, as demonstrated by the review conducted by Xiong et al. (2025).

Methodologically, data charting, extraction, and analysis were conducted by only one of the authors, despite recommendations for multiple coders to enhance reliability (Booth et al., 2022; Levac et al., 2010). This was mitigated in part by consulting an expert on suggestions for literature at key stages. The review also did not include forward citation searching. Although backward reference screening was conducted, future reviews should combine both approaches to improve coverage.

Additionally, Google Scholar screening was limited to the first five pages. While relevance declined after the first three pages, applying a fixed cutoff risks omitting relevant studies. Future reviews should consider more flexible stopping criteria based on relevance and saturation.

Second, the relatively small number of included studies ($n = 30$) is a limitation, resulting from the strict definition of data donation. This narrow scope was necessary, as data collection and sharing procedures are central to understanding participation. Given the limited literature, five studies with unclear procedures were included if they plausibly involved data donation and contributed to understanding willingness to donate.

Finally, most studies originate from a health context, which may limit generalizability. In health research, individuals tend to be more willing to donate data (e.g., Buhr et al., 2019) and may view it as a social duty (Skatova & Goulding,

2019). This may not apply to other contexts, such as information-seeking research (Blassnig et al., 2023), where willingness may differ. Findings should therefore be interpreted as context-dependent.

At the same time, the interdisciplinary nature of data donation research offers opportunities. Communication research in this area is still emerging and depends on recent developments in data access. Insights from other fields, where data sharing has been studied longer, can provide valuable input, encourage new perspectives, and support reflection on how findings apply to communication research.

Differences in definitions across disciplines require scrutiny to ensure a shared understanding, especially in cross-disciplinary collaborations. Establishing common terminology and methodological standards is essential for effective cooperation. When this is achieved, such collaborations can be highly productive. For example, social media data can be relevant not only to communication research but also to health research on mental well-being (Sleigh et al., 2018).

6. Conclusion

In conclusion, this study provides an overview of factors influencing willingness to donate and offers insights into strategies that may increase participation in data donation studies. It highlights participation bias in this context and presents actionable recommendations for study design.

To encourage participation, researchers should prioritize trust, transparency, and ethical standards such as data minimization. While not all studies can fully align with participant preferences, reducing effort and providing support remain important. Additionally, nudges and improved communication can be effective. In communication research, emphasizing societal relevance and individual benefits may further enhance engagement.

Considering these aspects can help make data donation studies more participant-centered, ethically sound, and impactful.

Nonetheless, beyond participation bias, other biases related to trace data must also be considered (Amaya et al., 2020; Keusch et al., 2023). Further research is needed to assess these limitations. Researchers should carefully select methods, samples, and measurements, particularly during study design, to minimize bias.

Overall, data donations offer strong potential for communication research by leveraging digital trace data to better understand societal phenomena and behavior, while helping address limitations of traditional methods. To fully realize this potential, future research should continue to examine participation motivations to improve methodology, mobilization, and representativeness.

Generative AI declaration

During the preparation of this manuscript, we used the following AI tools: DeepL Write was used for language editing and enhancing the clarity and flow of the drafted text. All AI-generated outputs were carefully reviewed and edited by the authors.

Disclosure statement

All authors declare that they have no conflicts of interest.

Funding

This research project is funded by the Bavarian Research Institute for Digital Transformation (bidt), an institute of the Bavarian Academy of Sciences and Humanities.

Supplementary material

Supplementary material for this article is available on the Open Science Framework (OSF): https://osf.io/rczpg/overview?view_only=c769a960b0914d30a3d16eb92872053e.

References

- Amaya, A., Biemer, P. P., & Kinyon, D. (2020). Total error in a big data world: Adapting the TSE framework to big data. *Journal of Survey Statistics and Methodology*, 8(1), 89–119. <https://doi.org/10.1093/jssam/smz056>
- Android.com. (2018, September 25). *Technology for today's world: Helping you reclaim a sense of balance*. Retrieved from <https://blog.google/products/android/technology-to-days-world-helping-you-reclaim-sense-balance/>
- Apple Inc. (2018, September 25). *iOS 12 führt neue Funktionen zum Vermeiden von Unterbrechungen und zur Verwaltung der Screen Time ein* [iOS 12 introduces new function to avoid interruptions and managing screen time]. Retrieved from <https://www.apple.com/de/newsroom/2018/06/ios-12-introduces-new-features-to-reduce-interruptions-and-manage-screen-time/>
- Aquisti, A. (2004, May 17). Privacy in electronic commerce and the economics of immediate gratification. *Proceedings of the 5th ACM Conference on Electronic Commerce*, New York, NY, USA. <https://doi.org/10.1145/988772.988777>
- Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods and Measures*, 11(3), 173–190. <https://doi.org/10.1080/19312458.2017.1317337>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Baumgartner, S. E., Sumter, S. R., Petkevič, V., & Wiradhany, W. (2023). A novel iOS data donation approach: Automatic processing, compliance, and reactivity in a longitudinal study. *Social Science Computer Review*, 41(4), 1456–1472. <https://doi.org/10.1177/08944393211071068>
- Bietz, M., Patrick, K., & Bloss, C. (2019). Data donation as a model for citizen science health research. *Citizen Science: Theory and Practice*, 4(1). <https://doi.org/10.5334/cstp.178>
- Blassnig, S., Mitova, E., Pfiffner, N., & Reiss, M. V. (2023). Googling referendum campaigns: Analyzing online search patterns regarding Swiss direct-democratic votes. *Media and Communication*, 11(1), 19–30. <https://doi.org/10.17645/mac.v11i1.6030>

- Boeschoten, L., Ausloos, J., Möller, J. E., Araujo, T., & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, 4(2), 388–423. <https://doi.org/10.5117/CCR2022.2.002.BOES>
- Booth, A., Sutton, A., Clowes, M., & Martyn-St James, M. (2022). *Systematic approaches to a successful literature review* (3rd ed.). SAGE.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Breuer, J., Kmetty, Z., Haim, M., & Stier, S. (2022). User-centric approaches for collecting Facebook data in the ‘post-API age’: Experiences from two studies and recommendations for future research. *Information, Communication & Society*, 26(14), 2649–2668. <https://doi.org/10.1080/1369118X.2022.2097015>
- Buhr, L., Schickntanz, S., & Nordmeyer, E. (2022). Attitudes toward mobile apps for pandemic research among smartphone users in Germany: National survey. *JMIR MHealth and UHealth*, 10(1), 1–17. <https://doi.org/10.2196/31857>
- Carrière, T. C., Boeschoten, L., Struminskaya, B., Janssen, H. L., Schipper, N. C. de, & Araujo, T. (2025). Best practices for studies using digital data donation. *Quality & Quantity*, 59(Suppl 1), 389–412. <https://doi.org/10.1007/s11135-024-01983-x>
- de León, E., Votta, F., Jongerius, S., Mulder, J., Struminskaya, B., Araujo, T., & de Vreese, C. (2025). Can tailored recruitment messaging increase digital trace data donation compliance? *International Journal of Social Research Methodology*, 1–17. <https://doi.org/10.1080/13645579.2025.2568497>
- Diethei, D., Niess, J., Stellmacher, C., Stefanidi, E., & Schöning, J. (2021). Sharing heartbeats: Motivations of citizen scientists in times of crises. In Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, & S. Drucker (Eds.), *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). ACM. <https://doi.org/10.1145/3411764.3445665>
- Dinev, T., McConnell, A. R., & Smith, H. J. (2015). Research commentary—Informing privacy research through information systems, psychology, and behavioral economics: Thinking outside the “APCO” box. *Information Systems Research*, 26(4), 639–655. <https://doi.org/10.1287/isre.2015.0600>
- European Parliament and the Council of the European Union. (2016, April 27). *On the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General data protection regulation)*. Retrieved from <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
- Gil-López, T., Christner, C., León, E. de, Makhortykh, M., Urman, A., Maier, M., & Adam, S. (2023). Do (not!) track me: Relationship between willingness to participate and sample composition in online information behavior tracking research. *Social Science Computer Review*, 41(6), 2274–2292. <https://doi.org/10.1177/08944393231156634>
- Gomez Ortega, A., Bourgeois, J., & Kortuem, G. (2021). Towards designerly data donation. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers* (pp. 496–501). ACM. <https://doi.org/10.1145/3460418.3479362#>
- Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 475–495. <https://doi.org/10.1086/269338>
- Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the filter bubble? *Digital Journalism*, 6(3), 330–343. <https://doi.org/10.1080/21670811.2017.1338145>

- Haim, M., & Hase, V. (2023). Computational Methods und Tools für die Erhebung und Auswertung von Social-Media-Daten [Computational methods and tools for social media data collection and analysis]. In S. Stollfuß, L. Niebling, & F. Raczkowski (Eds.), *Handbuch Digitale Medien und Methoden* (pp. 1–20). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36629-2_41-1
- Hakobyan, O., Hillmann, P.-J., Martin, F., Böttinger, E., & Drimalla, H. (2025). Development and evaluation of Dona, a privacy-preserving donation platform for messaging data from WhatsApp, Facebook, and Instagram. *Behavior Research Methods*, 57(3), 94. <https://doi.org/10.3758/s13428-024-02593-z>
- Hase, V., Ausloos, J., Boeschoten, L., Pfiffner, N., Janssen, H., Araujo, T., Carrière, T., de Vreese, C., Haßler, J., Loecherbach, F., Kmetty, Z., Möller, J., Ohme, J., Schmidbauer, E., Struminskaya, B., Trilling, D., Welbers, K., & Haim, M. (2024). Fulfilling data access obligations: How could (and should) platforms facilitate data donation studies? *Internet Policy Review*, 13(3). <https://doi.org/10.14763/2024.3.1793>
- Hase, V., & Haim, M. (2024). Can we get rid of bias? Mitigating systematic error in data donation studies through survey design strategies. *Computational Communication Research*, 6(2), 1–29. <https://doi.org/10.5117/CCR2024.2.2.HASE>
- Hendricks-Sturup, R., & Lu, C. Y. (2023). An assessment of perspectives and concerns among research participants of childbearing age regarding the health-relatedness of data, online data privacy, and donating data to researchers: Survey study. *Journal of Medical Internet Research*, 25. <https://doi.org/10.2196/41937>
- Hendricks-Sturup, R., Zhang, F., & Lu, C. Y. (2022). A survey of research participants' privacy-related experiences and willingness to share real-world data with researchers. *Journal of Personalized Medicine*, 12(11). <https://doi.org/10.3390/jpm12111922>
- Hillebrand, K., Hornuf, L., Müller, B., & Vrankar, D. (2023). The social dilemma of big data: Donating personal data to promote social welfare. *Information and Organization*, 33(1). <https://doi.org/10.1016/j.infoandorg.2023.100452>
- Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems*, 12(12), 767–797. <https://doi.org/10.17705/1jais.00282>
- Juga, J., Juntunen, J., & Koivumäki, T. (2021). Willingness to share personal health information: Impact of attitudes, trust and control. *Records Management Journal*, 31(1), 48–59. <https://doi.org/10.1108/RMJ-02-2020-0005>
- Karampela, M., Ouhbi, S., & Isomursu, M. (2019). Connected Health user willingness to share personal health data: Questionnaire study. *Journal of Medical Internet Research*, 21(11). <https://doi.org/10.2196/14537>
- Kaspar, K. (2020). Motivations COVID-19 for social distancing and app use as complementary measures to combat the pandemic: Quantitative survey study. *Journal of Medical Internet Research*, 22(8). <https://doi.org/10.2196/21613>
- Kellermeyer, L., Harnke, B., & Knight, S. (2018). Covidence and Rayyan. *Journal of the Medical Library Association*, 106(4). <https://doi.org/10.5195/jmla.2018.513>
- Keusch, F. (2015). Why do people participate in web surveys? Applying survey participation theory to internet survey data collection. *Management Review Quarterly*, 65(3), 183–216. <https://doi.org/10.1007/s11301-014-0111-y>
- Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., & Trappmann, M. (2023). Coverage error in data collection combining mobile surveys with passive measurement using apps: Data from a German national survey. *Sociological Methods & Research*, 52(2), 841–878. <https://doi.org/10.1177/00491241210914924>
- Keusch, F., Pankowska, P. K., Cernat, A., & Bach, R. L. (2024). Do you have two minutes to talk about your data? Willingness to participate and nonparticipation bias in Face-

- book data donation. *Field Methods*, 36(4), 279–293. <https://doi.org/10.1177/1525822X231225907>
- Keusch, F., Struminskaya, B., Mulder, J., & Jongerius, S. (2024). *Measuring physical activity in older adults through data donation: Consent rates, donation success, & bias*. <https://doi.org/10.31235/osf.io/fdg8t>
- Kmetty, Z., Stefkovics, Á., Számely, J., Deng, D., Kellner, A., Pauló, E., Omodei, E., & Kol-tai, J. (2024). Determinants of willingness to donate data from social media platforms. *Information, Communication & Society*, 28(7), 1324–1349. <https://doi.org/10.1080/1369118X.2024.2340995>
- Kulzer, B., Heinemann, L., & Roos, T. (2022). Patients' experience of new technologies and digitalization in diabetes care in Germany. *Journal of Diabetes Science and Technology*, 16(6), 1521–1531. <https://doi.org/10.1177/19322968211041377>
- Levac, D., Colquhoun, H., & O'Brien, K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5(69). <https://doi.org/10.1186/1748-5908-5-69>
- Marreiros, H., Tonin, M., Vlassopoulos, M., & Schraefel, M. C. (2017). “Now that you mention it”: A survey experiment on information, inattention and online privacy. *Journal of Economic Behavior & Organization*, 140, 1–17. <https://doi.org/10.1016/j.jebo.2017.03.024>
- Maus, B., Salvi, D., & Olsson, C. M. (2020). Enhancing citizens trust in technologies for data donation in clinical research: Validation of a design prototype. *Companion Proceedings of the 10th International Conference on the Internet of Things (IoT 2020)*. Advance online publication. <https://doi.org/10.1145/3423423.3423430>
- Mezinska, S., Kaleja, J., Mileiko, I., Santare, D., Rovite, V., & Tzivian, L. (2020). Public awareness of and attitudes towards research biobanks in Latvia. *BMC Medical Ethics*, 21(1). <https://doi.org/10.1186/s12910-020-00506-1>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1). <https://doi.org/10.1186/2046-4053-4-1>
- Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1). <https://doi.org/10.1186/s12874-018-0611-x>
- Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press. <https://doi.org/10.1515/9780804772891>
- Norberg, P., & Horne, D. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41(1), 100–126. <https://doi.org/10.1111/j.1745-6606.2006.00070.x>
- Ohme, J., & Araujo, T. (2022). Digital data donations: A quest for best practices. *Patterns*, 3(4). <https://doi.org/10.1016/j.patter.2022.100467>
- Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2023). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. *Communication Methods and Measures*, 18(2), 124–141. <https://doi.org/10.1080/19312458.2023.2181319>
- Ohme, J., Araujo, T., de Vreese, C., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, 9(2), 293–313. <https://doi.org/10.1177/2050157920959106>
- Pak, C., Cotter, K., & Thorson, K. (2022). Correcting sample selection bias of historical digital trace data: Inverse probability weighting (IPW) and type II Tobit model. *Com-*

- munication Methods and Measures*, 16(2), 134–155. <https://doi.org/10.1080/19312458.2022.2037537>
- Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis*, 18(10), 2119–2126. <https://doi.org/10.11124/JBIES-20-00167>
- Pfiffner, N., & Friemel, T. N. (2023). Leveraging data donations for communication research: Exploring drivers behind the willingness to donate. *Communication Methods and Measures*, 17(3), 227–249. <https://doi.org/10.1080/19312458.2023.2176474>
- Pilgrim, K., & Bohnet-Joschko, S. (2022). Effectiveness of digital forced-choice nudges for voluntary data donation by health self-trackers in Germany: Web-based experiment. *Journal of Medical Internet Research*, 24(2), 1–13. <https://doi.org/10.2196/31363>
- Prainsack, B. (2019). Data donation: How to resist the iLeviathan. In J. Krutzinna & L. Floridi (Eds.), *The ethics of medical data donation* (pp. 9–22). Springer. https://doi.org/10.1007/978-3-030-04363-6_2
- Scharkow, M. (2016). The accuracy of self-reported internet use—A validation study using client log data. *Communication Methods and Measures*, 10(1), 13–27. <https://doi.org/10.1080/19312458.2015.1118446>
- Seltzer, E., Goldshear, J., Guntuku, S. C., Grande, D., Asch, D. A., Klinger, E. V., & Merchant, R. M. (2019). Patients' willingness to share digital health and non-health data for research: A cross-sectional study. *BMC Medical Informatics and Decision Making*, 19(1). <https://doi.org/10.1186/s12911-019-0886-9>
- Severin-Nielsen, M. K. (2023). Politicians' social media usage in a hybrid media environment: A scoping review of the literature between 2008–2022. *Nordicom Review*, 44(2), 172–193. <https://doi.org/10.2478/nor-2023-0010>
- Silber, H., Breuer, J., Beuthner, C., Gummer, T., Keusch, F., Siegers, P., Stier, S., & Weiß, B. (2022). Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behaviour. *Journal of the Royal Statistical Society Series a: Statistics in Society*, 185(Supplement_2), S387–S407. <https://doi.org/10.1111/rssa.12954>
- Silber, H., Breuer, J., Felderer, B., Gerdon, F., Stammann, P., Daikeler, J., Keusch, F., & Weiß, B. (2024). Asking for traces: A vignette study on acceptability norms and personal willingness to donate digital trace data. *Social Science Computer Review*. Advance online publication. <https://doi.org/10.1177/08944393241305776>
- Skatova, A., & Goulding, J. (2019). Psychology of personal data donation. *PloS One*, 14(11), 1–20. <https://doi.org/10.1371/journal.pone.0224240>
- Sleigh, J. (2018). Experiences of donating personal data to mental health research: An explorative anthropological study. *Biomedical Informatics Insights*, 10. <https://doi.org/10.1177/1178222618785131>
- Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review*, 38(5), 503–516. <https://doi.org/10.1177/0894439319843669>
- Strycharz, J., Meppelink, C., Zarouali, B., Araujo, T., & Voorveld, H. (2024). The blind spot in data donations: Who is (not) willing to donate digital data in social scientific research. *Computational Communication Research*, 6(2), 1–27. <https://doi.org/10.5117/CCR2024.2.3.STRY>
- Susha, I., Grönlund, Å., & van Tulder, R. (2019). Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. *Government Information Quarterly*, 36(1), 112–128. <https://doi.org/10.1016/j.giq.2018.11.002>
- Thaler, R. H. (2018). Nudge, not sludge. *Science*, 361(6401), 431. <https://doi.org/10.1126/science.aau9241>

- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale Univ. Press.
- Toepoel, V., Luiten, A., & Zandvliet, R. (2021). Response, willingness, and data donation in a study on accelerometer possession in the general population. *Survey Practice*, 14(1), 1–19. <https://doi.org/10.29115/SP-2021-0005>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). Prisma extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- van Driel, I. I., Giachanou, A., Pouwels, J. L., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2022). Promises and pitfalls of social media data donations. *Communication Methods and Measures*, 16(4), 266–282. <https://doi.org/10.1080/19312458.2022.2109608>
- Wedel, L., Ohme, J., & Araujo, T. (2024). Augmenting data download packages – Integrating data donations, video metadata, and the multimodal nature of audio-visual content. Advance online publication. <https://doi.org/10.12758/mda.2024.08>
- Welbers, K., Loecherbach, F., Lin, Z., & Trilling, D. (2024). Anything you would like to share: Evaluating a data donation application in a survey and field study. *Computational Communication Research*, 6(2), 1–28. <https://doi.org/10.5117/CCR2024.2.5.WELB>
- Xiong, Y., van der Wal, A., & Beyens, I. (2025). Improving participation in data donation studies: A systematic review of factors driving participation and evidence-informed best practices. *Social Science Computer Review*. Advance online publication. <https://doi.org/10.1177/08944393251395958>
- Zannettou, S., Nemeth, O.-N., Ayalon, O., Goetzen, A., Gummadi, K. P., Redmiles, E. M., & Roesner, F. (2023, January 12). Analyzing user engagement with TikTok's short format video recommendations using data donations. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA (pp. 1–16). <https://doi.org/10.1145/3613904.3642433>

FULL PAPER

Beyond fact-checking: Harnessing AI for news quality assessment

Beyond Fact-Checking: KI-gestützte Qualitätsbewertung im Journalismus

Tobias Schmidt, Jonas Rieger, Jörg Rahnenführer, Astrid Viciano & Holger Wormer

Tobias Schmidt (M. A.), TU Dortmund University, Institute and School of Journalism, Emil-Figge-Straße 50, 44227 Dortmund, Germany. Contact: tobias3.schmidt@tu-dortmund.de. ORCID: <https://orcid.org/0000-0001-7330-4943>

Jonas Rieger (Dr.), TU Dortmund University, Department of Statistics, Vogelpothsweg 87, 44227 Dortmund, Germany. Contact: rieger@statistik.tu-dortmund.de. ORCID: <https://orcid.org/0000-0002-0007-4478>

Jörg Rahnenführer (Prof. Dr.), TU Dortmund University, Department of Statistics, Vogelpothsweg 87, 44227 Dortmund, Germany. Contact: rahnenfuehrer@statistik.tu-dortmund.de. ORCID: <https://orcid.org/0000-0002-8947-440X>

Astrid Viciano (Dr.), TU Dortmund University, Institute and School of Journalism, Emil-Figge-Straße 50, 44227 Dortmund, Germany. Contact: astrid.viciano@tu-dortmund.de.

Holger Wormer (Prof. Dipl.-Chem.), TU Dortmund University, Institute and School of Journalism, Emil-Figge-Straße 50, 44227 Dortmund, Germany. Contact: holger.wormer@tu-dortmund.de. ORCID: <https://orcid.org/0009-0005-7708-2934>



Beyond fact-checking: Harnessing AI for news quality assessment

Beyond Fact-Checking: KI-gestützte Qualitätsbewertung im Journalismus

Tobias Schmidt, Jonas Rieger, Jörg Rahnenführer, Astrid Viciano & Holger Wormer

Abstract: Reliable scientific information is crucial for shaping public opinion and informing economic, social, and health-related decisions. While significant research has focused on detecting misinformation, the automated assessment of the overall quality of journalistic content remains largely unexplored. Addressing this gap, our study investigates whether different artificial intelligence approaches ranging from classical statistical methods (“bag-of-words”) to advanced large language models (LLMs) can assist in evaluating the quality of medical news articles based on ten established journalistic quality criteria. Drawing on a dataset of 240 articles and corresponding expert reviews, we show that machine learning models can indeed assess article quality with considerable accuracy. While several models fall short of human-level performance, fine-tuned LLMs perform well and, in most cases, match the level of agreement with a consolidated expert benchmark observed among individual human reviewers. The best-performing model, a fine-tuned GPT-4o model, detects poor-quality articles reliably when evaluated against expert assessments (macro-averaged F1: 0.69 vs. 0.57), suggesting its potential for use in editorial workflows, such as filtering incoming material or double-checking an already written text. Although our study focuses on medical journalism’s well-defined criteria, the approach likely holds promise for enhancing quality assurance in other areas of journalism and science communication where similar quality criteria exist.

Keywords: Journalistic quality, science communication, machine learning, large language models, automation

Zusammenfassung: Zuverlässige wissenschaftliche Informationen sind entscheidend für die öffentliche Meinungsbildung sowie für wirtschaftliche, soziale und gesundheitspolitische Entscheidungen. Während sich zahlreiche Studien mit der Erkennung von Falschinformationen befassen, ist die automatisierte Bewertung der Gesamtqualität journalistischer Inhalte bislang weitgehend unerforscht. Diese Studie schließt diese Forschungslücke, indem sie untersucht, inwiefern verschiedene KI-Modelle – von klassischen statistischen Methoden („Bag-of-Words“) bis hin zu fortgeschrittenen Large Language Models (LLMs) – die Qualität medizinischer Nachrichtenartikel anhand von zehn etablierten journalistischen Qualitätskriterien bewerten können. Anhand eines Datensatzes mit 240 medizinjournalistischen Artikeln und zugehörigen Expertengutachten zu diesen Texten zeigen wir, dass Machine-Learning-Modelle durchaus in der Lage sind, die journalistische Qualität eines Artikels valide zu beurteilen. Zwar erreichen mehrere Modelle nicht das Leistungsniveau menschlicher Expert*innen, speziell feingetunte LLMs jedoch erzielen in den meisten Fällen ein

Maß an Übereinstimmung mit einer konsolidierten Expertenbewertung, das mit dem Übereinstimmungsgrad zwischen menschlichen Expert*innen vergleichbar ist. Das leistungsstärkste Modell (ein feingetunt GPT-4o) identifiziert qualitativ schlechte Artikel ebenso verlässlich wie menschliche Gutachter*innen (macro-averaged F1: 0.69 vs. 0.57) und deutet damit auf ein Potenzial für den Einsatz in redaktionellen Workflows hin, etwa zur Filterung eingehender Inhalte oder zur Nachkontrolle bereits verfasster Texte. Obwohl sich die Studie auf Qualitätskriterien aus dem Medizinjournalismus stützt, dürfte sich der Ansatz auch in anderen Bereichen des Journalismus und der Wissenschaftskommunikation anwenden lassen, sofern dort vergleichbare Qualitätskriterien existieren.

Schlagwörter: journalistische Qualität, Wissenschaftskommunikation, maschinelles Lernen, Large Language Models, Automatisierung

1. Introduction

Misinformation has become “something of a catch-all term” (Cacciatore, 2021, p. 1) and has attracted the attention of a broad range of disciplines. As early as 2020, a systematic review focused on news creation and consumption identified 2277 fake news-related scientific studies in the literature (Kim et al., 2021). The need for such studies is more important than ever, as we experience that “fake news is spreading day by day” (Jindal et al., 2023, p. 375). At the same time, several authors have pointed out that fostering trust in media may require more than simply combating misinformation in already published articles (e.g., Acerbi et al., 2022). Sorting out the worst misinformation and the most brazen lies does not automatically mean that the audience of (digital) media gets trustworthy and useful information, as even “seemingly trustworthy news sites push misleading headlines” (West & Bergstrom, 2021, p. 1). Following the idea that interventions should primarily seek to build trust in factual information (Tay et al., 2024), it becomes essential to ensure that information sources themselves consistently meet high-quality standards. In view of the cutbacks in many editorial offices, however, this is becoming an increasing challenge for the remaining staff.

Given these constraints, a growing number of studies have explored the potential of artificial intelligence (AI) tools to support journalistic workflows, especially in detecting and counteracting misinformation (Islam et al., 2020; Kumar et al., 2020). However, despite a growing recognition that a “fighting for information” could be far more effective than merely “fighting misinformation” (Acerbi et al., 2022), the potential of AI to assist with the systematic assessment of overall journalistic quality – beyond mere fact-checking – remains largely underexplored (Liu et al., 2025).

In this article, we therefore want to approach the semi-automated evaluation of news content from the other side of the scale between intentional disinformation (as probably the worst kind of information) and good (that is, trustworthy and useful) information. Using reporting on medical research as a case study, we assess the capability of several machine learning (ML) models, including large language models (LLMs), to evaluate the quality of science news content based on established criteria for good medical journalism. The example of medical research as a scientific topic seems particularly suitable for two reasons: First, re-

porting on medical sciences not only accounts constantly for the largest share of science reporting (Bucchi & Mazzolini, 2003; Clark & Illman, 2006; Summ & Volpers, 2016) but misleading information in this field has also the largest potential of causing harm among media recipients (Schwitzer, 2008). Secondly, the quality criteria for good reporting (partly based on standards known from evidence-based medicine) are probably the most developed and precise in science journalism, if not in journalism at all. The proposed concept as well as the developed models presented in this study could, however, be extended to other kinds of reporting, too, given that an elaborated set of quality criteria and standards has been developed.

In the following, we first provide an overview of established concepts and quality criteria for “good reporting” in general, and for science communication in particular, with a specific focus on medical journalism. Next, we examine recent research on ML applications in journalism and explore the potential of AI tools to contribute to quality assessment beyond fact-checking alone. We then propose a methodological framework for developing and evaluating AI tools specifically designed for quality assessment in journalism. Specifically, we evaluate the performance of two common statistical classification models, several LLMs used in a zero-shot setting – that is, without any task-specific fine-tuning – and three fine-tuned LLMs in autonomously assessing the quality of science news articles based on predefined criteria. By comparing these AI-driven assessments with human expert evaluations, we identify the strengths and limitations of each model. Finally, we discuss the adaptability of our approach and suggest future directions for integrating AI-driven quality assessment into the evolving digital media landscape.

2. Related work

Quality of public information is a broad issue which has been discussed in theory and practice ever since public media existed. Numerous terms have emerged to conceptualise this field, with Denis McQuail’s (1993) “media performance” being among the most frequently cited concepts. Other commonly used terms include “journalistic” or “newspaper excellence” (Gladney, 1996), while the phrase “quality in journalism” is often preferred in German-language journalism research. Some frequently mentioned general criteria for quality in the media include *facticity* and *accuracy*, *relevance*, *timeliness*, *independence*, *diversity of topics and sources*, and *usefulness*, among others (Meier, 2019). Many of these criteria are applicable beyond traditional journalism, extending to other communication formats like public relations or social media posts (Meier, 2019) and are also mentioned by science communication researchers (Fähnrich et al., 2023; Silva Luna et al., 2025). However, catalogues and recommendations for communicating specific scientific disciplines to the public often go beyond the general quality criteria mentioned above.

One of the first criteria-based approaches to measure the content quality of reports from a specific scientific discipline can be found in Canada (Oxman et al., 1993). With their index of scientific quality (ISQ) for health reports in the lay

press, Oxman et al. (1993) built on earlier work in the field of the so-called accuracy research, concentrating rather on lists of error categories than on quality criteria (Tankard & Ryan, 1974). A systematic monitoring of the quality of medical news reporting using defined criteria combined with a peer-review system came from the *Media Doctor* project in Australia (Smith et al., 2005; Wilson et al., 2009), followed by *HealthNewsReview.org* in the US with the same ten criteria (Schwitzer, 2008).

Building on this international tradition, the German quality assessment initiative “Medien-Doktor” was introduced in 2010, following the same conceptual approach (Wormer, 2011). It has since been adapted to other communication formats (e.g., press releases) and further developed for additional scientific domains, such as environmental reporting (Rögener & Wormer, 2017). Today, it constitutes the largest publicly available repository of expert-reviewed science news articles in the German context, offering a unique resource for empirical research on science journalism quality. Given its systematic methodology, transparent evaluation scheme, and rich corpus of annotated news articles, the Medien-Doktor framework is particularly well-suited for the present study’s aim to benchmark automated models against human expert evaluations. We therefore adopt its ten medicine-related quality criteria as the foundation for our analysis, as summarized in Table 1.

Table 1. Overview of the ten Medien-Doktor quality criteria, adapted and reordered from Schwitzer (2008) and the former HealthNewsReview.org in the US.

Criterion 1	Article adequately explains benefits of a new drug/treatment/procedure...
Criterion 2	Adequately explains potential harms
Criterion 3	Reviews the study methodology or the quality of the evidence
Criterion 4	Seeks out independent sources and discloses potential conflicts of interest
Criterion 5	Appears not to rely solely on a news release
Criterion 6	Establishes the true novelty of the idea
Criterion 7	Compares the new idea with existing alternatives
Criterion 8	Establishes the availability of the product or procedure
Criterion 9	Adequately discusses costs
Criterion 10	Avoids disease mongering

These ten criteria aim to make quality in medical journalism assessable in a structured, transparent manner. Ideally, a high-quality article would fulfil all ten criteria, by clearly stating the benefits (criterion 1) and risks (2) of a new procedure, explaining the evidence base (3), providing information on availability (8), communicating all of this in a sober and factual tone (10), and so forth. In the time-pressured reality of journalistic practice, however, this ideal is rarely achieved. The Medien-Doktor project, much like its international counterparts, therefore, pursues a dual goal: To sensitise journalists to these criteria and to make transpa-

rent for the public which articles merit closer scrutiny. For a detailed rationale and the development process of the criteria, see Wormer (2011) and Schwitzer (2008).

2.1 AI systems in the newsroom

Established quality criteria like the ones presented above hold the potential to assist journalists and science communicators in their everyday tasks. They facilitate the screening of incoming press releases or agency news, help identify weaknesses and low-quality documents and support the detection of the most promising material, which can then serve as a starting point for thorough investigations. At the same time, it is widely acknowledged that most newsrooms face significant time and labour constraints. Screening articles and potentially double-checking already written texts takes time, a resource often missing in modern newsrooms. As a result, time pressure not only leads to less diversity and overall, less cross-checking of articles (Reich & Godler, 2014), but also impacts communicators' ability to effectively fact-check information (Himma-Kadakas & Ojamets, 2022). Some argue that, under such circumstances, journalism is prone to "become churnalism, the recycling of pre-packaged public relations and press agency copy" (Davies, 2008, p. 59) – a challenge that persists even in the presence of well-defined quality criteria.

Given these constraints and external factors such as competitive dynamics and pressure to innovate, media outlets have increasingly turned to automation tools designed to support journalists in their daily work (Hermida & Simon, 2025; Simon, 2024). Commonly referred to as part of *Automated journalism* (Carlson, 2015; Caswell & Dörr, 2018; Diakopoulos, 2019), *Algorithmic journalism* (Dörr, 2016; Kotenidis & Veglis, 2021), or *Computational journalism* (Cohen et al., 2011; Thurman, 2019), these tools typically involve the use of algorithms and data analysis to support journalistic tasks. This can include identifying newsworthy trends or stories within data (Liu et al., 2016; Stray, 2019), producing routine articles on topics like sports or finance, or structuring news websites based on recommender systems (Bodó et al., 2019; Latar, 2018). Other tools assist with analyzing unstructured documents, data wrangling or lead generation. Especially the potential of ML models to detect and counteract misinformation has aroused attention in this context (Islam et al., 2020; Kumar et al., 2020).

With the advent of generative artificial intelligence (GenAI) and particularly LLMs, the implementation of AI tools in science communication in general (Schäfer, 2023) and especially in journalism has gained new momentum (Diakopoulos et al., 2024; Simon, 2024). In this context, the term *journalistic Artificial Intelligence* or *journalistic AI* has gained prominence, often used as an umbrella term for all kinds of technologies, "from machine learning to automated decision making, that are used along the entire news production chain" (Helberger et al., 2022, p. 1606). Modern systems are used to summarize or rephrase texts, to generate headlines, personalize news, or assist in real-time fact-checking and background research. Compared to earlier tools, their accessibility and versatility have made them far more present in everyday journalistic practice. Correspondingly,

expectations have grown that such tools could relieve journalists from repetitive and time-consuming tasks, allowing them to dedicate more time to investigative reporting and complex research (Van Dalen, 2012; Wolf, 2024).

However, while there is considerable enthusiasm surrounding the use of modern journalistic AI in newsrooms, the task of evaluating the overall quality of content is rarely mentioned in this context. When discussing the potential of AI systems to filter articles, e.g., researchers typically elaborate on verification, or claim-matching, rather than systems to check the overall article quality (Simon, 2024). As discussed earlier, this approach might be insufficient for effectively promoting trustworthy information. Perhaps most strikingly, the Journalism AI project from the London School of Economics and Political Science lists numerous case studies from around the world on how AI systems are currently utilised in journalism – and reveals that none of the > 250 presented case studies (as of March 2026, subject to future updates) employ AI tools for the quality assessment of news content (LSE, 2025). Most of the related research and practical applications predominantly focus on news production and misinformation detection. That is, the extent to which AI systems can evaluate or “double-check” the overall quality of media content beyond fact-checking remains unclear.

A rare example of studies discussing automation tools for evaluating (science) news quality rather than detecting fake news was presented by Løvlie et al. (2023). The authors developed and evaluated a “Scientific Evidence Indicator” (SEI) designed to help readers assess the strength of evidence and uncertainty in science journalism articles. Their tool evaluates scientific studies based on measurable criteria such as peer review status and methodological quality. Another related study was put forward by Choi et al. (2021), who trained an artificial neural network (ANN) from scratch to predict audience-related news quality assessment. Combining traditional survey methods with ANN-training and other computational approaches, the authors identified journalistic values such as believability, depth, and diversity as stronger predictors of audience-rated news quality than linguistic features or audience attributes. The study that comes closest to our approach was performed by Liu et al. (2025). The authors used the expert-annotated *HealthNewsReview* dataset to evaluate the capability of traditional ML models and *GPT-3.5-Turbo* to rate the quality of health news across established journalistic criteria. Their findings show that while *GPT-3.5* can generate clear and contextually relevant explanations, its actual rating accuracy remains limited and falls short of the traditional ML models. Thus, the authors demonstrate both the potential and the shortcomings of one specific LLM when applied to multi-dimensional quality evaluation tasks, leaving room for further research in this direction. Beyond these few studies, however, the potential of ML models to support criteria-based quality evaluation of journalistic content remains largely unexplored.

2.2 LLMs in content classification: Promise, pitfalls, and alternatives

Large language models, in particular modern state-of-the-art (SOTA) models like OpenAI's *GPT-5* or Anthropic's *Claude 4.5 Sonnet* have demonstrated strong performance across a variety of natural language processing (NLP) tasks, including natural language understanding, text classification, and information extraction (Chang et al., 2023; Chiang et al., 2024). Especially their ability to understand and process highly diverse texts and their ease of access position them as an effective tool for evaluating diverse types of articles when no other assistant systems exist. Given these strengths, it may seem intuitive to regard LLMs as a default solution for tasks like the one addressed in this study.

At the same time, the use of LLMs is not without concern. Several researchers highlight risks associated with uncritical adoption of AI (and especially LLM) outputs (Mahony & Chen, 2025; Wu, 2024). The “black-box” nature of LLMs makes it difficult to fully understand the decision-making process behind their outputs, which can lead to ethical and practical challenges in journalism and even provoke distrust and hostility (Bail, 2024; Noain Sánchez, 2022; Schäfer et al., 2024). Furthermore, generative AI tools may display various biases, not only producing offensive content, such as racist or misogynistic language, but also hallucinating inaccurate information that could further exacerbate the already existing challenges of disseminating trustworthy information (Schäfer, 2023). Many researchers even recognize threats to editorial independence and the potential erosion of core journalistic values such as accuracy, transparency, and accountability (Cordero et al., 2025; Møller et al., 2025; Wolf, 2024). Last but not least, the increasing reliance on proprietary LLMs – often developed by a small number of global tech companies – can reinforce structural dependencies and limit the autonomy of newsrooms (Simon, 2024).

Considering these challenges, exploring more transparent and interpretable modeling approaches becomes increasingly important. Targeted ML solutions, such as those designed for specific classification tasks, may offer more reliability and control than general-purpose, free-text-generating systems, especially in classification tasks. Classical statistical methods such as Random Forests (Breiman, 2001) or Naive Bayes classifiers (McCallum & Nigam, 1998) offer greater potential for interpretability, which can be crucial in contexts where transparency and accountability are required. Furthermore, several studies have shown that LLMs do not consistently outperform simpler models across all tasks (Liu et al., 2025; Qin et al., 2023; Ziemis et al., 2024). Thus, we see considerable value in not solely evaluating *ChatGPT*'s capabilities at a certain task (as is often the case in communication science research) but also benchmarking it against other model types.

2.3 Objective and research questions

Following these concerns, our research setup includes both classical statistical models and several state-of-the-art LLMs. This complementary approach makes it possible to investigate not only the capabilities of general-purpose language models but also the potential of simpler, more transparent classification systems

for content evaluation (Liu et al., 2025). Additionally, we incorporate fine-tuning of LLMs, utilising both small open-source and large proprietary models. This setup allows us to identify the strengths and limitations of each model – and to test whether the limitations reported by Liu et al. (2025) persist regardless of the choice of model and once models are adapted to domain-specific data.

We ask the following research questions: Using the example of medicine-related news articles, can machine learning models reliably assess article quality? Precisely, how far can these models approach human expert performance in evaluating the content quality based on established quality criteria? (RQ1) What are the minimum technical requirements (in terms of model size and complexity) needed to achieve expert-level accuracy? (RQ2) How does prediction accuracy vary across different quality criteria? (RQ3) To what extent can machine learning models detect articles of low quality, serving as inbound or outbound control? (RQ4) Following these questions, we aim to provide insights into how, and to what extent, AI can assist journalists in assessing the overall quality of various texts, be it an incoming press release, material from a news agency, or their own articles, as a final check before publication.

3. Data and methods

We evaluate the effectiveness of a Random Forest (RF) classifier, a Naive Bayes (NB) model, the zero-shot (“out-of-the-box”) capabilities of several LLMs, and the performance of three different LLMs that were fine-tuned (ft) specifically on the present task.

Random Forest and Naive Bayes are two well-established classification algorithms commonly used in text analysis. Naive Bayes is a probabilistic model based on Bayes’ theorem that assumes independence between features/words, an assumption that is often violated in natural language but still delivers robust results for classification tasks (Zhang, 2004). Random Forest, in contrast, is an ensemble learning method that combines the predictions of multiple decision trees to improve classification accuracy and control overfitting (Breiman, 2001). Both methods rely on simple word-frequency-based representations of text (commonly referred to as “bag-of-words” assumption), making them particularly suitable for settings that prioritise interpretability, resource efficiency, and methodological transparency.¹ In addition, their simplicity makes inference computationally much less intensive, making them suitable for rapid analysis or applications where computational resources are limited.

As mentioned earlier, LLMs provide a fundamentally different approach to text evaluation, as they operate on semantic representations rather than surface-level word frequencies. While the NB and RF models in our implementation operate on lexical features, LLMs can leverage contextual understanding and reason-

¹ It is important to note that “transparency” in this context typically denotes the potential to examine how model features contribute to predictions, for instance by tracing influential words or analysing interactions between them. Still, performing such analyses requires additional effort and does not necessarily translate into valuable practical insights.

ning to evaluate complex criteria in unstructured text. Furthermore, LLMs are particularly advantageous when annotated data are scarce – which are needed for training a NB or RF model –, as they are pretrained and can simply be instructed in natural language. However, given the findings of Liu et al. (2025), we expect moderate performance levels from these models and hypothesise that the specialized ML models exceed the zero-shot capabilities of certain LLMs, especially older and smaller ones. To evaluate the capabilities of LLMs, we test a variety of models that are chosen for their strong performance on various tasks, as indicated by the Chatbot Arena LLM Leaderboard (Chiang et al., 2024). Our selection includes both small and large models, as well as proprietary and open-source models.

Additionally, we perform LLM fine-tuning, since several researchers demonstrate LLM’s potential in coding/annotation tasks when specifically trained (Gilardi et al., 2023; Mellon et al., 2024; Rieger et al., 2024). Fine-tuning makes it possible to adapt models to a specific domain (medical-related media articles in our case) and tailor their responses, which significantly improves model performance, especially in areas where precision and task-specific knowledge are crucial (Wu et al., 2023). Precisely, we test the open-source model *Mistral-7B-Instruct-v0.3* from Mistral AI, which we access via the open-source model library Hugging Face, and two models offered by OpenAI, *GPT-4o* and *GPT-4o-mini*.

An overview of all tested models is presented in Table 2. Supplementary material for this article is available on the Open Science Framework (OSF): <https://osf.io/6yk82/>.

Table 2. Overview of all tested models.

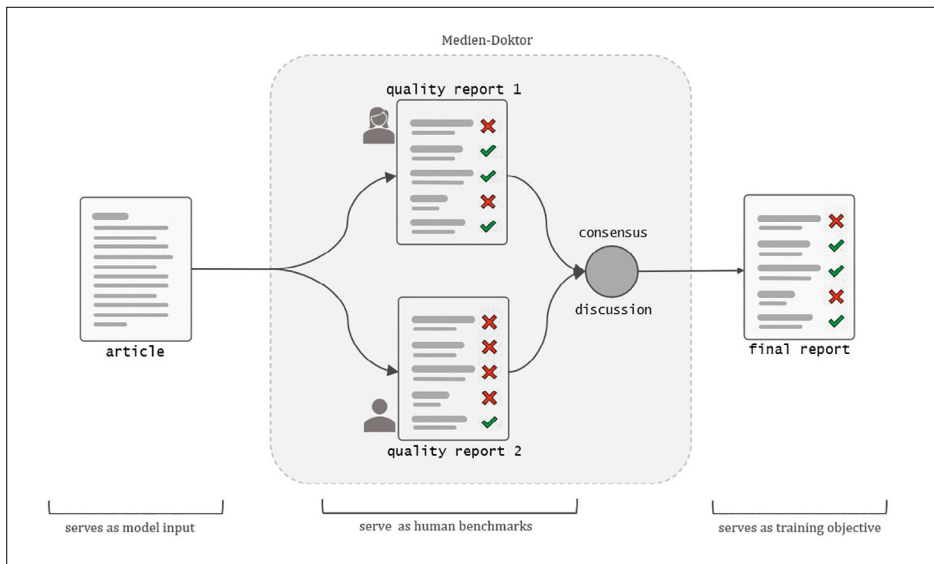
Bag-of-words models	Large language models	
	Zero-shot LLM	Fine-tuned LLM
Naive Bayes	GPT-3.5-turbo, GPT-4o, GPT-4-1106-Preview, Claude-3.5-Sonnet	GPT-4o-mini
Random Forest	Mistral 7B, Mistral 8x22B, Mistral Large	GPT-4o
	Llama 3 8B, Llama 3 70B	Mistral-7B-Instruct-v0.3

3.1 Training setup

As a benchmark for all tests, we resort to the human quality evaluations corresponding to the science news articles we assess ($N = 240$). Both articles and reports have been published by the Medien-Doktor project, which offers the largest publicly available repository of medicine-related news reviews in a German context. The quality reports are consolidated summaries based on two independent evaluations for each article. For each article evaluation, the Medien-Doktor commissions two science journalism experts (“peers”) specialised in the relevant field. These experts independently assess the texts based on the criteria presented above and then confer to agree on a consolidated overall evaluation. The evaluations include both a label (criterion *fulfilled*, *not fulfilled*, *not applicable*) and a brief

explanation of the decision taken for each criterion. As it is also the case in other peer review processes, this approach inherently introduces some level of disagreement between individual expert evaluations and the final summary ratings, providing a benchmark for reasonable evaluation discrepancies. For each article, the expert report that shows higher agreement with the consolidated summary is referred to as *human_high* in the following, while the remaining report is referred to as *human_low*. The terms “high” and “low” refer exclusively to the level of agreement with the consolidated expert summary and do not indicate differences in assessment quality. This joint report is treated as gold standard as it reflects the consensus reached by the editor between two reviewers. This procedure is illustrated in Figure 1.

Figure 1. Illustration of the evaluation procedure used in the Medien-Doktor project



Notes. Each article is independently reviewed by two domain-specialized science journalism experts, who produce individual quality reports based on the ten predefined criteria (*quality reports 1 & 2*). The reviewers then discuss their assessments and agree on a consolidated *final report*, which serves as the gold standard in our study. For each article, the individual report that aligns more closely with the final consensus is labelled *human_high* (in this example: Reviewer 1), while the one showing a greater deviation is labelled *human_low* (Reviewer 2).

3.2 Naive Bayes models

For our Naive Bayes approach, we train ten separate models, each specialised in predicting the assessment of one of the quality criteria by returning a single label $\mathcal{E} \{ \text{fulfilled, not fulfilled} \}$. This one-model-per-criterion setup is necessary since each quality criterion represents an independent binary classification task, and NB does not jointly model multiple labels in a single classifier. The models are

trained on a document feature matrix representing our training texts. We include only case-sensitive features (unigrams) that occur at least four times in the whole training data and in at least four documents. A small set of stop words (mainly pronouns and articles) and numbers is removed. This yields a medium-sized vocabulary that captures general lexical regularities without overfitting to idiosyncratic expressions in single articles. The final configuration was chosen based on a systematic comparison of 128 NB variants with different parameter combinations and n-gram settings based on our training data.

3.3 Random Forest classifier

In addition to this generic bag-of-words baseline, we train ten Random Forests, one for each of the ten quality criteria. In contrast to the NB models, the RFs operate on a compact, expert-informed feature set: For every criterion, the authors compiled a list of domain-specific signal words and regular expressions based on the established quality criteria (see Table 1) and their domain expertise in science journalism (e.g., terms referring to study design for “evidence” or to adverse effects for “risks”; see S1 b.). For each article and each criterion, we count how often these criterion-specific patterns occur and use these counts as the only features in the corresponding RF. Thus, each RF is trained on a small, criterion-tailored feature vector whose dimensionality equals the number of signal words defined for that criterion. This design deliberately differs from the NB setup: While NB serves as a classical, broad bag-of-words baseline, the RF models encode expert knowledge in a focused “bag-of-cues” representation, allowing us to test whether curated signal words alone can approximate human quality ratings. In the following, we refer to NB as well as the RF approaches as bag-of-words or “BOW” models in the sense that they disregard word order and rely exclusively on token or pattern counts.

3.4 Zero-Shot LLMs

When using an LLM “zero-shot”, we give each model, one by one, an article as input, accompanied by the ten quality criteria presented earlier. All models are prompted to behave like a “rigorous scientific reviewer of health journalism articles”, thereby following the well-established practice of assigning a specific role to an LLM to enhance its performance (Salewski et al., 2024; Shanahan et al., 2023). In each prompt, we include all criteria simultaneously along with the article, requiring the model to respond to all criteria sequentially as *fulfilled*, *not fulfilled* or *not applicable*. We also experimented with different prompting strategies, the differences in performance, however, remained minor. Specifically, our results showed that instructing the models to generate a full textual review before assigning labels slightly reduced accuracy and did not provide additional analytical value in the zero-shot setting (see S1 c. for further details on the prompting). In the following, when referring to a language model that is used in this vein, we use the term “zero-shot model”.

3.5 Fine-tuned LLMs

Our “ft”-models are trained to generate standardised quality reports in natural language in the style of a human-written review, as the ones provided by the Medien-Doktor project (see S1 d. for an illustration). Specifically, the models process a raw document as input and output a detailed report consisting of ten paragraphs, each beginning with a brief assessment and concluding with a label $\mathcal{E}$ [fulfilled, not fulfilled, not applicable]. The idea behind this approach is that, during training, the models internalise how the criteria are applied in context by “reading” expert evaluations and producing their own versions, with the fine-tuning process penalising deviations from the human assessments. This is a notable difference from the zero-shot LLMs, that apply the criteria based solely on fixed definitions they receive as part of the prompt.

We train all models on the same 212 documents of the human reviewers that were sampled randomly from the Medien-Doktor database. The remaining 28 documents serve as test data. The sampling ensures that the test set includes a representative distribution of article types and quality levels. Importantly, none of the models has access to the test articles during training, and no overlapping or near-duplicate content exists between the sets. The findings presented below are based on the models’ (and humans’) quality evaluations of the 28 test documents. For the evaluation, all predictions, including the LLM outputs and the expert classifications, are converted into a binary format by treating all “not applicable” labels as “not fulfilled” (see S1 a., supplementary material on OSF, for further details). Since each report evaluates the same ten criteria, the total number of data points for analysis is 280 (10 criteria $\times$ 28 documents).

To contextualise the dataset, Table 3 summarises the distribution of fulfilled and non-fulfilled labels across the 28 test documents. In addition, to give readers a sense of the types of articles analysed, four randomly selected headlines (English translations) are listed below (Table 4).

Table 3. Overview of label distribution across all ten criteria (C1–C10) in the test set. Values indicate the proportion of articles for which a given criterion was rated as fulfilled or not fulfilled

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
fulfilled	0.43	0.54	0.39	0.36	0.68	0.68	0.29	0.79	0.50	0.71
not fulfilled	0.57	0.46	0.61	0.64	0.32	0.32	0.71	0.21	0.50	0.29

Notes. C1 = Adequately explains benefits; C2 = Adequately explains potential harms; C3 = Reviews the study methodology or quality of evidence; C4 = Seeks independent sources and discloses conflicts of interest; C5 = Does not rely solely on a news release; C6 = Establishes the true novelty of the idea; C7 = Compares the new idea with existing alternatives; C8 = Establishes availability of the product or procedure; C9 = Adequately discusses costs; C10 = Avoids disease mongering.

Table 4. Excerpt of the dataset to be evaluated

Source	Date	Title (translated by the authors)
RND	2019-06-04	New coffee study: Up to 25 cups a day are safe
Der Standard	2020-02-17	Pregnancy: Overweight due to parabens in cosmetics
T-Online	2019-09-30	Cheese protects blood vessels from damage caused by excessive salt intake
Tagesspiegel	2020-08-04	Does Russia already have the coronavirus vaccine?

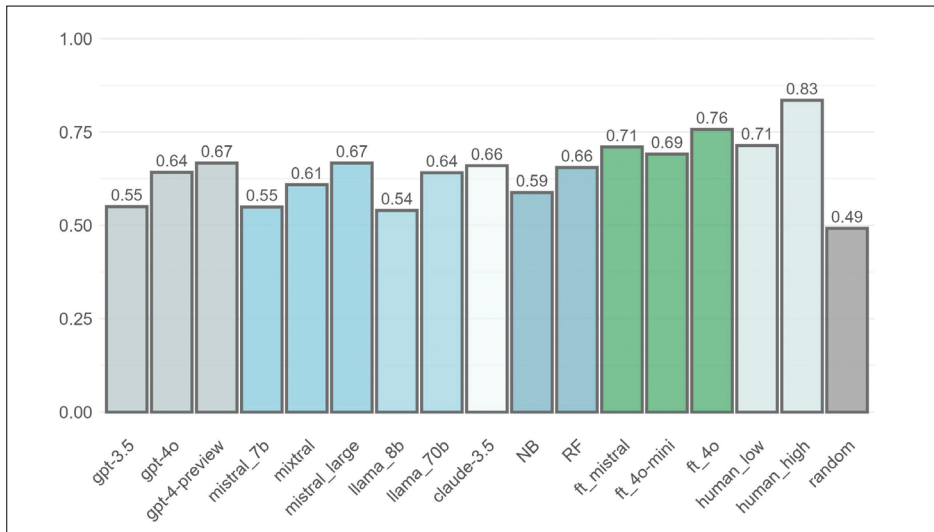
4. Results

Based on the examined case of medical reporting, our findings suggest that ML models may offer useful support for assessing the quality of journalistic content. Particularly, fine-tuned LLMs appear to hold promise in assisting journalists with this task. However, some important differences remain, which may be relevant for the design of potential future assistance systems.

4.1 Overall model performances

Many tested models achieve reasonable accuracy rates in aligning their criterion evaluations with the consolidated expert assessments. For example, the best zero-shot LLMs, *GPT-4-1106-Preview* and *Mistral Large*, correctly predict 187 out of 280 labels, achieving accuracy rates of 67%, which is four percentage points below the values of the “*human_low* reports” (198 labels). However, since class imbalances affect our dataset, accuracy rates alone are insufficient to evaluate performance. Therefore, we additionally assess precision, recall, and macro-averaged F1-scores for both labels (“fulfilled” and “not fulfilled”). These numbers confirm that larger zero-shot LLMs come close to human expert results, and significantly outperform smaller and older models (e.g., *GPT-3.5-Turbo*, *Mistral 7B*, *Llama 8B*). Noteworthy, the RF classifier is among the best-performing models, closely matching the accuracy of the top-performing zero-shot LLMs (Figure 2). As described earlier, these rates should not be interpreted against certain reliability rates of a coding process (e.g., intercoder reliability) but against the background of current agreement rates in a peer-review process, where rather moderate levels of agreement among reviewers in similar contexts are common (Bornmann et al., 2010; Burla et al., 2008; Reinhart, 2009).

The highest performance is achieved by the fine-tuned LLMs, all of which reach F1-scores of 0.69 or higher and even surpass our “*human_low*” baseline (0.71). When comparing the zero-shot *Mistral 7B* to its fine-tuned counterpart, the performance gains from fine-tuning become particularly evident (0.55 vs. 0.71 F1 score).

Figure 2. Macro-averaged F1 scores achieved by all tested models

Notes. The different colour sets indicate the different model families: *OpenAI* (grey), *Mistral* (blue), *Llama* (light blue), *Claude* (white), BOW models (dark blue), fine-tuned LLMs (green; “ft_” indicates fine-tuning), human experts (light grey). For comparison, the figure also displays the performances of a hypothetical model that would sample randomly from both labels (“random”). Values reflect agreement with the subjective evaluations by (other) experts rather than with an objective ground truth. As in other peer-review processes, moderate levels of interrater agreement are expected.

Beyond quantitative performance, a heuristic examination of the quality reports generated by the fine-tuned models provides further insights into the models’ reasoning processes and their alignment with human evaluations. We find that fine-tuning indeed taught the models to focus on similar aspects as human experts, such as prioritizing absolute over relative numbers when assessing an article’s description of the benefits (criterion 1) or risks (criterion 2) of a treatment (see S4, supplementary material on OSF for an example). For the smaller two ft-models, the results are not as clear-cut. While performing well in quantitative terms, a heuristic analysis of the generated reports reveals that the smaller models partly learned to overfit the training data. In an illustrative example, the *fine-tuned Mistral* model learned to always classify criterion 10 as fulfilled, stating that “[topic of article] is not exaggerated”. This clearly shows the importance of model selection and parameter tuning, a challenge we want to tackle further in future research.

These data let us answer research questions 1–2 as follows: Regarding RQ1, our results demonstrate that automated quality assessment of scientific news articles is indeed feasible, with fine-tuned LLMs achieving F1 scores of up to 0.76 – a performance that surpasses that of human experts. Given both the performance metrics and the substantive quality of the model-generated reports, the *fine-tuned GPT-4o* appears capable of producing assessments as reliable as those written by the lower-performing half of the human experts.

As for RQ2, concerning the minimum technical requirements needed for reliable quality assessments, our findings show that while simple NB models and older LLMs like *GPT-3.5-Turbo* perform only marginally better than random guessing, an RF classifier can match the performance of more complex LLMs. We conclude that while using more recent *ChatGPT* variants with detailed prompts (which is a common procedure in many newsrooms) provides a convenient approach with reasonable accuracy, similar outcomes can be achieved with a simpler RF model. This confirms earlier findings by Liu et al. (2025) and underlines the importance of not focusing on popular LLMs like *ChatGPT* alone (which is often the case in public discourse), but also considering simpler, more interpretable models where appropriate. For optimal performance, however, fine-tuning an LLM appears essential.

4.2 Criterion-specific performance

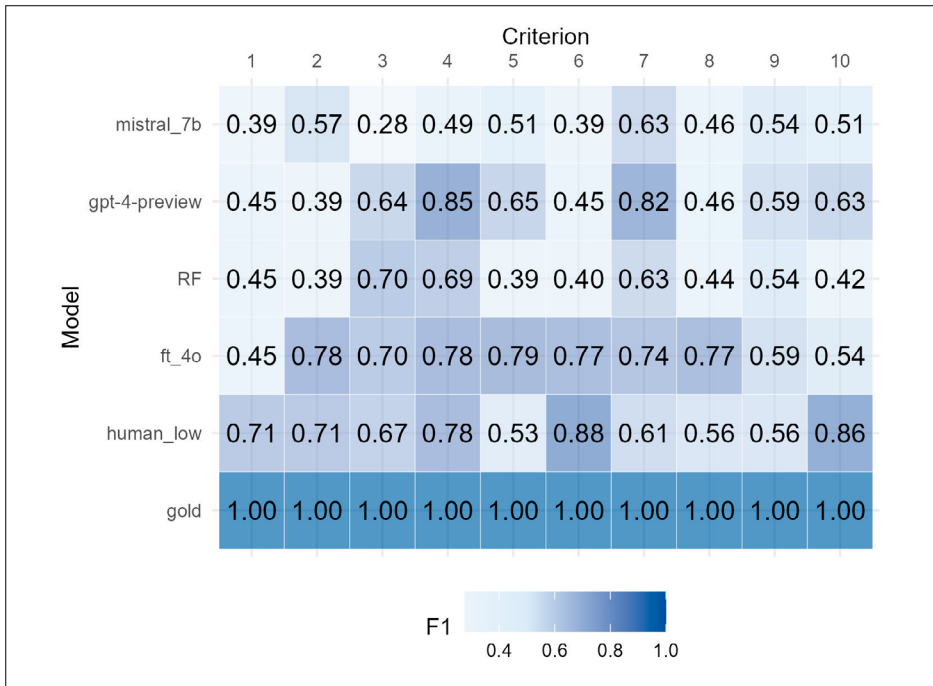
To better understand the strengths and weaknesses of the tested models, we evaluate their performance on each of the ten criteria individually (S3 b.). The results show that, while the prediction of certain criteria appears straightforward, multiple criteria remarkably challenge the tested models (Figure 3). Especially the criteria 1 (potential benefits) and 10 (neutrality/avoiding disease mongering) prove difficult to classify, with all models frequently diverging from human expert conclusions (min. 0.2 difference in avg. F1 score, respectively). While the fine-tuned language models achieve comparatively strong results, they still struggle with these two criteria.

A likely explanation for this lies in the underlying training data distribution. For certain criteria, especially criterion 10, most gold-standard labels are “fulfilled” (20 out of 28). As a result, the ft-models appear to have learned this dominant pattern well, yielding high recall and precision for positive predictions, but struggle to reliably identify rare “not fulfilled” cases. This issue might be mitigated in future through targeted adjustments to the training data, such as data augmentation (see Discussion). However, such adjustments will not be a universal fix. For example, criterion 9 (costs) also poses difficulties, yet here even human experts show relatively poor agreement, suggesting that some criteria may be inherently harder to assess. In the context of journalistic practice, these findings once again demonstrate the importance of understanding a model’s “behaviour” when relying on its outputs. In our case, this would mean that predictions for criteria 1 and 10 would require special attention, with human judgement assuming a particularly prominent role in their evaluation.

In sum, regarding RQ3, our findings reveal substantial variation in prediction accuracy across quality criteria. While some criteria are predicted with reasonable recall and precision, multiple consistently prove challenging for both BOW and the zero-shot LLMs. Only the fine-tuned LLMs, especially the *fine-tuned GPT-4o* model, manage to reliably predict eight out of ten criteria, demonstrating the benefits of training on complete assessment reports. However, the persistent difficulty with specific criteria suggests that even fine-tuned models may struggle to gene-

realize across all dimensions of article quality, pointing to structural limitations that warrant further investigation.

Figure 3. Criterion-wise model performance



Notes. The heatmap shows macro-averaged F1 scores for each of the ten quality criteria (columns) across a selection of models (rows). Included are: *Mistral 7B* (representing smaller open-source LLMs), *GPT-4-1106-Preview* (as the best-performing zero-shot LLM), Random Forest (RF; best-performing bag-of-words model), and a *fine-tuned GPT-4o* model (*ft_4o*). Human performance is represented by the less-aligned reports (*human_low*) and the expert consensus (*gold*). The figure highlights that while certain criteria are reliably predicted by most models, others, such as criteria 1 and 10, remain challenging even for fine-tuned LLMs.

4.3 Quality-based article selection

In practical applications, the overall assessment of an article is arguably more important than a model's accuracy on individual criteria. To address this, we also evaluate the models' ability to perform a "gatekeeping" function by distinguishing between high- and low-quality articles, based on a document's number of fulfilled criteria. To this end, we categorise each article based on its expert rating: Articles with fewer than five fulfilled criteria are considered "low quality," those with only one or two fulfilled criteria are classified as "bad". All other articles are classified as "pass", i.e., articles that fulfil five or more of the ten criteria.

Our experiments show that all top-performing models (RF, *gpt-4-preview*, *ft_4o*) successfully classify all the worst articles ($n = 5$) as either low-quality or

bad (see S5 for an illustration). Given the small number of worst articles, this result should be interpreted as illustrative rather than statistically conclusive. Overall, the best-performing model (*fine-tuned GPT-4o*) correctly classifies 21 out of 28 articles, achieving higher agreement with the consolidated expert assessment than the less-aligned expert reports (*human_low*, 19/28). The other models, in contrast, show substantial variation in their class prediction capabilities, with, for example, the zero-shot *GPT-4o* model misclassifying two of the best articles as low-quality. This illustrates two things: First, the benefits of fine-tuning an LLM on expert data, and second, the inherent complexity and subjectivity of the task, as even human experts do not always fully agree on which articles should be considered low quality. Table 5 summarises the results in the form of a confusion matrix (recall, precision, F1).

Table 5. Macro-averaged precision, recall, and F1-scores across the three quality classes (bad, low quality, pass) for selected models and the human expert baseline

	precision	recall	F1
gpt-4-preview	0.46	0.44	0.43
RF	0.42	0.53	0.47
ft_4o	0.68	0.73	0.69
human_low	0.57	0.56	0.57

Considering this, we conclude (RQ4) that various types of models can provide meaningful assistance in editorial workflows, particularly during filtering stages. Again, the *fine-tuned GPT-4o* model delivered the most reliable results. However, none of these models should be seen as a substitute for expert review. They should instead be understood as complementary tools that support, rather than replace, human judgement.

5. Discussion

The results of our study indicate that the evaluation and potential adoption of AI models for content evaluation beyond fact-checking is a promising direction for both journalism and science communication research. We show that clearly defined quality criteria, particularly in a scientific domain, can be used for ML-based content-evaluation of articles, given these models are either properly trained or large enough to be able to perform an analysis zero-shot. For incoming texts such as press releases from scientific institutions, agency materials or pieces from freelancers, a reliably performing model could provide initial quality checks and identify articles that appear most promising for further (human) investigation. Before the final release of journalistic texts, it could furthermore serve as an outbound control, ensuring that content meets established quality criteria of a certain media house or institution. However, among the tested models, only the fine-tuned LLMs, particularly the *fine-tuned GPT-4o*, reached a level of performance that makes them credible candidates for potential application in newsroom environ-

ments. Smaller and older models such as *GPT-3.5* performed significantly worse, especially compared to traditional ML models. This finding complements the results presented by Liu et al. (2025), where *GPT-3.5-Turbo* also showed limited accuracy in a zero-shot setting and underperformed relative to supervised baselines. Our results indicate that this outcome is largely attributable to the model's limited capabilities rather than a fundamental inability of LLMs to evaluate journalistic quality. More recent and larger models substantially outperform *GPT-3.5* in our tests and reach performance levels comparable to our Random Forest classifier. Crucially, we find that fine-tuning LLMs on full-text expert reports leads to further improvements, enabling the best-performing model to achieve higher agreement with the consolidated expert assessments than individual human reviewers. Furthermore, it comes with the benefit of providing users with a detailed report that discusses an article's quality regarding each of the ten quality criteria, which can be a valuable support in journalistic workflows.

At the same time, when it comes to the real-world adaptability of our framework, we acknowledge that fine-tuning a language model is not a trivial task, and that newsrooms might face challenges setting up their own quality-evaluation models. While ML expertise may play a role, the primary obstacle probably remains the lack of high-quality training data, which critically constrains a model's ability to learn nuanced quality distinctions (see below). With this in mind, there is a need for further investigation into how to further push the zero-shot performance of LLMs. Comprehensive analyses of where and why discrepancies between model and human evaluations occur could lead to more targeted improvements in model prompting. We are confident that through extensive prompt engineering, the performance of the zero-shot models analysed here can be further improved. However, whenever sufficiently high-quality training data is available, fine-tuning should be the preferred approach.

For any real-world application, it remains essential to emphasise that even the most capable model cannot and should not replace journalistic gatekeeping responsibilities. While models can assist in initial screening, they cannot and should not fully replace human oversight and critical judgement, which remain core pillars of the journalistic profession (Gutiérrez Lopez et al., 2023; Lin & Lewis, 2022). This is especially true for nuanced editorial decisions, when model outputs are ambiguous or when a model has already demonstrated systematic limitations in certain settings. For example, our results have shown that even a fine-tuned LLM that achieves higher agreement with the consolidated expert assessments than individual human reviewers in many tests shows systematic biases in predicting certain quality criteria. Encouragingly, recent studies confirm that most journalists, even when actively using AI tools to enhance efficiency, maintain a critical stance, insisting on verifying and editing AI outputs to ensure alignment with journalistic standards (Wu, 2024).

6. Limitations

Despite the promising results of our study, several limitations must be acknowledged. Firstly, our analysis focuses exclusively on the example of articles related to medical science and relies on the existence and testability of predefined quality criteria. Thus, the adaptability of our approach to other areas of journalism remains to be explored. We are, however, confident that given the existence of clear quality standards, our approach could be effectively applied in other fields. In particular, the utility of these models for economic or environmental reporting (where corresponding quality criteria have already been developed (Rögener & Wormer, 2017)) seems worth exploring in future research. Secondly, although we tested a variety of models and prompt techniques, the generalisability of our findings could be further strengthened by incorporating other popular (large language) models. In particular, resorting to more recent “thinking models” like *DeepSeek R1* or OpenAI’s *o3* model could be promising and should be evaluated in future research. The same applies to the potential of including tool-use or internet search capabilities in the LLM setup. Both extensions could allow LLMs to access external information, potentially enhancing their performance on tasks requiring up-to-date or context-specific knowledge. Since long-term independence from proprietary model providers such as OpenAI or Alphabet might be a relevant strategic consideration for many news organisations, open-source models such as *Mistral*, *Gemma 3* (Google) or *Phi 4* (Microsoft) deserve particular attention in this context. Many of these models can now be re-trained or adapted with comparatively modest computational resources, including setups based on a cloud platform such as Google Colab or Kaggle, indicating that, from a technical standpoint, the prerequisites for adopting and customising open-source solutions are already quite favourable. Finally, our findings represent a snapshot in time. Given the rapid pace of development in large language models, future architectures or model versions may exhibit different performance characteristics, even on comparable tasks.

A further critical limitation is the restricted size of our dataset. We are well aware that incomplete or limited training data contributes to a lack of robustness and prediction accuracy. Moreover, as with any fine-tuning on relatively small datasets, overfitting remains a potential concern – although our best-performing models showed no indications of overfitting on the validation or unseen test data (a brief analysis of the training dynamics is provided in supplementary material S1 d on OSF). However, the limited availability of publicly accessible expert reviews on science news articles, such as those presented here, constrains our ability to scale up the training data. One potential strategy to mitigate data scarcity is data augmentation (e.g., Feng et al., 2021; Wei & Zou, 2019). However, implementing such techniques in our context is challenging: Both the original articles and their corresponding quality assessments would need to be adapted in a highly controlled and semantically consistent manner. As an alternative or complementary strategy, collaborations with media organizations could offer a more sustainable path forward. By systematically logging editorial decisions over time (e.g., which texts are selected, edited, or discarded, and why so), these organizations could generate domain-specific training data that reflects real-world workflows and decisions.

7. Conclusions

Budget constraints in media organisations, coupled with an increasing need for efficient quality assessment amid growing information flows, have intensified the demand for tools that can support rapid and reliable decision-making. This is especially true in fields such as science reporting, where specific editorial knowledge is often rare. Addressing this need, our study combines theoretical insights from quality research with practical model evaluations, testing a variety of AI approaches for their potential to automate criteria-based quality assessment of news articles, particularly in the domain of medical science reporting. This approach moves beyond the established, albeit important, practice of fact-checking specific claims (Hameleers, 2019; Nyhan et al., 2020) towards a more holistic evaluation of journalistic content based on predefined criteria for the overall quality of a given contribution.

Rather than relying solely on the “black box” of LLMs, we conducted a comparative analysis to evaluate the capabilities of both LLMs and more transparent methods in terms of their ability to evaluate an article’s quality. By testing a range of models against established quality criteria for medical journalism, we found that current AI technologies can deliver results that are, in many cases, comparable to human expert evaluations. Both advanced LLMs, such as *GPT-4o*, and simpler models like a Random Forest classifier achieved accuracy levels approaching those of human evaluators but showed imbalances in their prediction capabilities of individual quality criteria. Notably, the fine-tuned LLMs employed in this study showed a significant advantage over their zero-shot counterparts. A *fine-tuned GPT-4o* model consistently aligned more closely with – and even surpassed – human expert evaluations. Nevertheless, at least two quality criteria proved challenging also for the fine-tuned models, underscoring the complexity and nuanced understanding needed for robust quality assessment in journalism. Importantly, the fine-tuned models also demonstrated strong capabilities in detecting low-quality articles, suggesting practical applications as assistive tools in editorial workflows, particularly during early screening stages. While our study focused on medical journalism, we are confident that our approach can be extended to other fields where clear quality standards exist or can be developed, tailored to the needs of specific organisations or aligned with theories of media performance. Ultimately, the models tested here should always be seen as augmenting, rather than replacing, human judgment.

Generative AI declaration

During the preparation of this work, the authors used DeepL and, in rare cases, ChatGPT to refine the linguistic quality and improve the readability of the manuscript. The authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Supplementary material

Supplementary material for this article is available on the Open Science Framework (OSF): <https://osf.io/6yk82>.

References

- Acerbi, A., Altay, S., & Mercier, H. (2022). Research note: Fighting misinformation or fighting for information? *Harvard Kennedy School (HKS) Misinformation Review*, 3, 1–15.
- Bail, C. A. (2024). Can generative AI improve social science? *Proceedings of the National Academy of Sciences*, 121(21). <https://doi.org/10.1073/pnas.2314021121>
- Bodó, B., Helberger, N., Eskens, S., & Möller, J. (2019). Interested in diversity: The role of user attitudes, algorithmic feedback loops, and policy in news personalization. *Digital journalism*, 7(2), 206–229. <https://doi.org/10.1080/21670811.2018.1521292>
- Bornmann, L., Mutz, R., & Daniel, H.-D. (2010). A reliability-generalization study of journal peer reviews: A multilevel meta-analysis of inter-rater reliability and its determinants. *PLoS ONE*, 5(12). <https://doi.org/10.1371/journal.pone.0014331>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bucchi, M., & Mazzolini, R. G. (2003). Big science, little news: Science coverage in the Italian daily press, 1946–1997. *Public Understanding of Science*, 12(1), 7–24. <https://doi.org/10.1177/0963662503012001413>
- Burla, L., Knierim, B., Barth, J., Liewald, K., Duetz, M., & Abel, T. (2008). From text to codings: Intercoder reliability assessment in qualitative content analysis. *Nursing Research*, 57(2), 113–117. <https://doi.org/10.1097/01.NNR.0000313482.33917.7d>
- Cacciatore, M. A. (2021). Misinformation and public opinion of science and health: Approaches, findings, and future directions. *Proceedings of the National Academy of Sciences*, 118(15). <https://doi.org/10.1073/pnas.1912437117>
- Carlson, M. (2015). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism*, 3(3), 416–431. <https://doi.org/10.1080/21670811.2014.976412>
- Caswell, D., & Dörr, K. (2018). Automated journalism 2.0: Event-driven narratives: From simple descriptions to real stories. *Journalism Practice*, 12(4), 477–496. <https://doi.org/10.1080/17512786.2017.1320773>
- Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P.S., Yang, Q., & Xie, X. (2023). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15, 1–45. <https://doi.org/10.1145/3641289>
- Chiang, W.-L., Zheng, L., Sheng, Y., Angelopoulos, A. N., Li, T., Li, D., Zhang, H., Zhu, B., Jordan, M., Gonzalez, J. E., & Stoica I. (2024). Chatbot arena: An open platform for evaluating LLMs by human preference. *arXiv*, arXiv:2403.04132. <https://doi.org/10.48550/arXiv.2403.04132>
- Choi, S., Shin, H., & Kang, S.-S. (2021). Predicting audience-rated news quality: Using survey, text mining, and neural network methods. *Digital Journalism*, 9(1), 84–105. <https://doi.org/10.1080/21670811.2020.1842777>
- Clark, F., & Illman, D. L. (2006). A longitudinal study of the New York Times science times section. *Science Communication*, 27(4), 496–513. <https://doi.org/10.1177/1075547006288010>

- Cohen, S., Hamilton, J. T., & Turner, F. (2011). Computational journalism. *Communications of the ACM*, 54(10), 66–71. <https://doi.org/10.1145/2001269.2001288>
- Cordero, J. M., Henn, T., Holtel, F., Sánchez Gómez, J. Á., Arenas, D., Šipka, A., & Vollmer, S. (2025). Developing effective and value-aligned AI tools for journalists: 12 critical questions to reflect upon. *Journalism Practice*, 1–20. <https://doi.org/10.1080/17512786.2025.2465894>
- Davies, N. (2008). *Flat earth news: An award-winning reporter exposes falsehood, distortion and propaganda in the global media*. Chatto & Windus.
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press. <https://doi.org/10.4159/9780674239302>
- Diakopoulos, N., Cools, H., Helberger, N., Li, C., Kung, E., & Rinehart, A. (2024). *Generative AI in journalism: The evolution of newswork and ethics in a generative information ecosystem*. Associated Press. <https://doi.org/10.13140/RG.2.2.31540.05765>
- Dörr, K. N. (2016). Mapping the field of algorithmic journalism. *Digital Journalism*, 4(6), 700–722. <https://doi.org/10.1080/21670811.2015.1096748>
- Fähnrich, B., Weitkamp, E., & Kupper, J. F. (2023). Exploring ‘quality’ in science communication online: Expert thoughts on how to assess and promote science communication quality in digital media contexts. *Public Understanding of Science*, 32(5), 605–621. <https://doi.org/10.1177/09636625221148054>
- Feng, S. Y., Gangal, V., Wei, J., Chandar, S., Vosoughi, S., Mitamura, T., & Hovy, E. (2021). A survey of data augmentation approaches for NLP. *arXiv preprint arXiv:2105.03075*. <https://doi.org/10.18653/v1/2021.findings-acl.84>
- Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30). <https://doi.org/10.1073/pnas.2305016120>
- Gladney, G. A. (1996). How editors and readers rank and rate the importance of eighteen traditional standards of newspaper excellence. *Journalism & Mass Communication Quarterly*, 73(2), 319–331. <https://doi.org/10.1177/107769909607300204>
- Gutiérrez Lopez, M., Porlezza, C., Cooper, G., Makri, S., MacFarlane, A., & Missaoui, S. (2023). A question of design: Strategies for embedding AI-driven tools into journalistic work routines. *Digital Journalism*, 11(3), 484–503. <https://doi.org/10.1080/21670811.2022.2043759>
- Hameleers, M. (2019). Susceptibility to mis- and disinformation and the effectiveness of fact-checkers: Can misinformation be effectively combated? *SCM Studies in Communication and Media*, 8(4), 523–546. <https://doi.org/10.5771/2192-4007-2019-4-523>
- Helberger, N., van Drunen, M., Moeller, J., Vrijenhoek, S., & Eskens, S. (2022). Towards a normative perspective on journalistic AI: Embracing the messy reality of normative ideals. *Digital Journalism*, 10(10), 1605–1626. <https://doi.org/10.1080/21670811.2022.2152195>
- Hermida, A., & Simon, F. M. (2025). AI in the newsroom: Lessons from the adoption of The Globe and Mail’s Sophi. *Journalism Practice*, 19(10), 1–18. <https://doi.org/10.1080/17512786.2025.2471781>
- Himma-Kadakas, M., & Ojamets, I. (2022). Debunking false information: Investigating journalists’ fact-checking skills. *Digital Journalism*, 10(5), 866–887. <https://doi.org/10.1080/21670811.2022.2043173>
- Islam, M. R., Liu, S., Wang, X., & Xu, G. (2020). Deep learning for misinformation detection on online social networks: A survey and new perspectives. *Social Network Analysis and Mining*, 10(1). <https://doi.org/10.1007/s13278-020-00696-x>

- Jindal, H., Mangla, M., & Singh, G. (2023, June 16–17). *Fake news detection using machine learning*. International Conference on Recent Developments in Cyber Security (pp. 375–385), Greater Noida, India. <https://doi.org/10.1016/j.procs.2025.09.048>
- Kim, B., Xiong, A., Lee, D., & Han, K. (2021). A systematic review on fake news research through the lens of news creation and consumption: Research efforts, challenges, and future directions. *PLoS ONE*, 16(12). <https://doi.org/10.1371/journal.pone.0260080>
- Kotenidis, E., & Veglis, A. (2021). Algorithmic journalism – current applications and future perspectives. *Journalism and Media*, 2(2), 244–257. <https://doi.org/10.3390/journalmedia2020014>
- Kumar, S., Asthana, R., Upadhyay, S., Upreti, N., & Akbar, M. (2020). Fake news detection using deep learning models: A novel approach. *Transactions on Emerging Telecommunications Technologies*, 31(2). <https://doi.org/10.1002/ett.3767>
- Latar, N. L. (2018). *Robot journalism: Can human journalism survive?*. World Scientific. <https://doi.org/10.1142/10913>
- Lin, B., & Lewis, S. C. (2022). The one thing journalistic AI just might do for democracy. *Digital Journalism*, 10(10), 1627–1649. <https://doi.org/10.1080/21670811.2022.2084131>
- Liu, X., He, L., Alanazi, E., Liu, E., Goss, A., & Gumireddy, L. (2025). Assessing the accuracy and explainability of using ChatGPT to evaluate the quality of health news. *BMC Public Health*, 25(1). <https://doi.org/10.1186/s12889-025-23206-0>
- Liu, X., Li, Q., Nourbakhsh, A., Fang, R., Thomas, M., Anderson, K., ..., & Shah, S. (2016). Reuters tracer: A large scale system of detecting & verifying real-time news events from Twitter. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management* (pp. 207–216). Association for Computing Machinery. <https://doi.org/10.1145/2983323.2983363>
- Løvlie, A. S., Waagstein, A., & Hyldgård, P. (2023). “How trustworthy is this research?” Designing a tool to help readers understand evidence and uncertainty in science journalism. *Digital Journalism*, 11(3), 431–464. <https://doi.org/10.1080/21670811.2023.2193344>
- LSE (London School of Economics and Political Science) (2025). *JournalismAI Case Studies*. Retrieved March 13, 2026, from <https://airtable.com/appdaeDwFizD4RK0u/shrKhe7Js48HvBhmG/tblBcSZESOAuy5Q9A>
- Mahony, S., & Chen, Q. (2025). Concerns about the role of artificial intelligence in journalism, and media manipulation. *Journalism*, 26(9). <https://doi.org/10.1177/14648849241263293>
- McCallum, A., & Nigam, K. (1998). A comparison of event models for naive Bayes text classification. In *AAAI-98 Workshop on Learning for Text Categorization* (pp. 41–48). AAAI Press.
- McQuail, D. (1993). *Media performance: Mass communication and the public interest*. University of Toronto Press. <https://doi.org/10.22230/cjc.1993v18n4a783>
- Meier, K. (2019). Quality in journalism. In T. P. Vos, F. Hanusch, D. Dimitrakopoulou, M. Geertsema-Sligh, & A. Sehl (Eds.), *The International Encyclopedia of Journalism Studies*. Wiley. <https://doi.org/10.1002/9781118841570.iej0041>
- Mellon, J., Bailey, J., Scott, R., Breckwoldt, J., Miori, M., & Schmedeman, P. (2024). Do AIs know what the most important issue is? Using language models to code open-text social survey responses at scale. *Research & Politics*, 11(1). <https://doi.org/10.1177/20531680241231468>
- Møller, L. A., van Dalen, A., & Skovsgaard, M. (2025). A little of that human touch: How regular journalists redefine their expertise in the face of artificial intelligence. *Journalism Studies*, 26(1), 84–100. <https://doi.org/10.1080/1461670X.2024.2412212>

- Noain Sánchez, A. (2022). Addressing the impact of artificial intelligence on journalism: The perception of experts, journalists and academics. *Communication & Society*, 35(3), 105–121. <https://doi.org/10.15581/003.35.3.105-121>
- Nyhan, B., Porter, E., Reifler, J., & Wood, T. J. (2020). Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior*, 42(3), 939–960. <https://psycnet.apa.org/doi/10.1007/s11109-019-09528-x>
- Oxman, A. D., Guyatt, G. H., Cook, D. J., Jaeschke, R., Heddle, N., & Keller, J. (1993). An index of scientific quality for health reports in the lay press. *Journal of Clinical Epidemiology*, 46, 987–1001. [https://doi.org/10.1016/0895-4356\(93\)90166-x](https://doi.org/10.1016/0895-4356(93)90166-x)
- Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M., & Yang, D. (2023). Is ChatGPT a general-purpose natural language processing task solver? *arXiv*, arXiv:2302.06476. <https://doi.org/10.18653/v1/2023.emnlp-main.85>
- Reich, Z., & Godler, Y. (2014). A time of uncertainty: The effects of reporters' time schedule on their work. *Journalism Studies*, 15(5), 607–618. <https://doi.org/10.1080/1461670X.2014.882484>
- Reinhart, M. (2009). Peer review of grant applications in biology and medicine. Reliability, fairness, and validity. *Scientometrics*, 81(3), 789–809. <https://doi.org/10.1007/s11192-008-2220-7>
- Rieger, J., Yanchenko, K., Ruckdeschel, M., von Nordheim, G., Kleinen-von Königsłow, K., & Wiedemann, G. (2024). Few-shot learning for automated content analysis: Efficient coding of arguments and claims in the debate on arms deliveries to Ukraine. *SCM Studies in Communication and Media*, 13(1), 72–100. <https://doi.org/10.5771/2192-4007-2024-1-72>
- Rögener, W., & Wormer, H. (2017). Defining criteria for good environmental journalism and testing their applicability: An environmental news review as a first step to more evidence based environmental science reporting. *Public Understanding of Science*, 26(4), 418–433. <https://doi.org/10.1177/0963662515597195>
- Salewski, L., Alaniz, S., Rio-Torto, I., Schulz, E., & Akata, Z. (2024). In-context impersonation reveals large language models' strengths and biases. *Advances in Neural Information Processing Systems*, 36, 72044–72057. <https://doi.org/10.48550/arXiv.2305.14930>
- Schäfer, M. S. (2023). The notorious GPT: Science communication in the age of artificial intelligence. *JCOM: Journal of Science Communication*, 22. <https://doi.org/10.22323/2.22020402>
- Schäfer, M. S., Kremer, B., Mede, N. G., & Fischer, L. (2024). Trust in science, trust in ChatGPT? How Germans think about generative AI as a source in science communication. *Journal of Science Communication*, 23(9). <https://doi.org/10.22323/2.23090204>
- Schwitzer, G. (2008). How do US journalists cover treatments, tests, products, and procedures? An evaluation of 500 stories. *PLoS Medicine*, 5(5). <https://doi.org/10.1371/journal.pmed.0050095>
- Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*, 623(7987), 493–498. <https://doi.org/10.1038/s41586-023-06647-8>
- Silva Luna, D., Broer, I., Bilandzic, H., Taddicken, M., Schuller, B. W., & Bürger, M. (2025). Quality in science communication with communicative artificial intelligence: A principle-based framework. *Public Understanding of Science*, 34(8). <https://doi.org/10.1177/09636625251328854>
- Simon, F. M. (2024). *Artificial intelligence in the news: How AI retools, rationalizes, and reshapes journalism and the public arena (tech. rep.)*. Tow Center for Digital Journalism, Columbia University. <https://doi.org/10.7916/ncm5-3v06>

- Smith, D. E., Wilson, A. J., & Henry, D. A. (2005). Monitoring the quality of medical news reporting: Early experience with media doctor. *Medical Journal of Australia*, 183(4), 190–193. <https://doi.org/10.5694/j.1326-5377.2005.tb06992.x>
- Stray, J. (2019). Making artificial intelligence work for investigative journalism. *Digital Journalism*, 7(8), 1076–1097. <https://doi.org/10.1080/21670811.2019.1630289>
- Summ, A., & Volpers, A.-M. (2016). What's science? Where's science? Science journalism in German print media. *Public Understanding of Science*, 25(7), 775–790. <https://doi.org/10.1177/0963662515583419>
- Tankard, J. W., Jr., & Ryan, M. (1974). News source perceptions of accuracy of science coverage. *Journalism Quarterly*, 51(2), 219–225. <https://doi.org/10.1177/107769907405100204>
- Tay, L. Q., Lewandowsky, S., Hurlstone, M. J., Kurz, T., & Ecker, U. K. (2024). Thinking clearly about misinformation. *Communications Psychology*, 2(1). <https://doi.org/10.1038/s44271-023-00054-5>
- Thurman, N. (2019). Computational journalism. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The Handbook of Journalism Studies* (2nd ed., pp. 180–195). Routledge.
- Van Dalen, A. (2012). The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists. *Journalism Practice*, 6(5–6), 648–658. <https://doi.org/10.1080/17512786.2012.667268>
- Wei, J., & Zou, K. (2019). EDA: Easy data augmentation techniques for boosting performance on text classification tasks. *arXiv preprint arXiv:1901.11196*. <https://doi.org/10.18653/v1/D19-1670>
- West, J. D., & Bergstrom, C. T. (2021). Misinformation in and about science. *Proceedings of the National Academy of Sciences*, 118(15). <https://doi.org/10.1073/pnas.1912444117>
- Wilson, A., Bonevski, B., Jones, A., & Henry, D. (2009). Media reporting of health interventions: Signs of improvement, but major problems persist. *PLoS ONE*, 4(3). <https://doi.org/10.1371/journal.pone.0004831>
- Wolf, C. (2024). Journalismus und künstliche Intelligenz aus kommunikationswissenschaftlicher Perspektive: Chancen und Herausforderungen [Journalism and artificial intelligence from a communication science perspective: Opportunities and challenges]. In G. Hooffacker, W. Kenntemich, & U. Kulisch (Eds.), *Neue Plattformen – neue Öffentlichkeiten: KI, Krisen und Journalismus* (pp. 9–29). Springer Fachmedien. https://doi.org/10.1007/978-3-658-44659-8_2
- Wormer, H. (2011). Improving health care journalism. In G. Gigerenzer & J. A. Muir Gray (Eds.), *Better doctors, better patients, better decisions: Envisioning Health Care 2020* (pp. 169–188). MIT Press. <https://doi.org/10.7551/mitpress/9143.003.0016>
- Wu, S. (2024). Journalists as individual users of artificial intelligence: Examining journalists' "value-motivated use" of ChatGPT and other AI tools within and without the newsroom. *Journalism*. <https://doi.org/10.1177/14648849241303047>
- Wu, S., Irsoy, O., Lu, S., Dabrowski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). BloombergGPT: A large language model for finance. *arXiv*, arXiv:2303.17564. <https://doi.org/10.48550/arXiv.2303.17564>
- Zhang, H. (2004). The optimality of naive Bayes. In V. Barr & Z. Markov (Eds.), *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (FLAIRS 2004)* (pp. 562–567). AAAI Press.
- Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2024). Can large language models transform computational social science? *Computational Linguistics*, 50, 237–291. https://doi.org/10.1162/coli_a_00502

FULL PAPER

Freie Rede auf dem Campus? Einflüsse von politischer Kongruenz und Sanktionsängsten auf die Meinungsfreiheitswahrnehmung und Redebereitschaft deutscher Studierender

Free speech on campus? The impact of political congruence and fear of sanctions on perceptions of freedom of speech and willingness to speak out among German university students

Luna Laufer, Franziska Selzer, Julia Stockwald, Annika Schleithoff, Nathalie L. Neuendorf, Anna-Luisa Sacher & Carsten Reinemann

Luna Laufer (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland.

Franziska Selzer (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland. Kontakt: franziska.selzer@ifkw.lmu.de. ORCID: <https://orcid.org/0009-0009-7572-0989>

Julia Stockwald (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland.

Annika Schleithoff, Deutsche Journalistenschule München, Hultschiner Str. 8, 81677 München, Deutschland.

Nathalie L. Neuendorf (M. A.), Universität Augsburg, Institut für Medien, Wissen und Kommunikation, Universitätsstr. 10, 86159 Augsburg, Deutschland. Kontakt: nathalie.neuendorf@uni-a.de. ORCID: <https://orcid.org/0000-0002-5938-8578>

Anna-Luisa Sacher (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland. Kontakt: anna.sacher@ifkw.lmu.de. ORCID: <https://orcid.org/0000-0002-3386-4310>

Carsten Reinemann (Prof. Dr.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland. Kontakt: carsten.reinemann@ifkw.lmu.de. ORCID: <https://orcid.org/0000-0001-9715-8434>



Freie Rede auf dem Campus? Einflüsse von politischer Kongruenz und Sanktionsängsten auf die Meinungsfreiheitswahrnehmung und Redebereitschaft deutscher Studierender

Free speech on campus? The impact of political congruence and fear of sanctions on perceptions of freedom of speech and willingness to speak out among German university students

Luna Laufer, Franziska Selzer, Julia Stockwald, Annika Schleithoff, Nathalie L. Neuendorf, Anna-Luisa Sacher & Carsten Reinemann

Zusammenfassung: Meinungsfreiheit ist nicht nur ein grundlegendes Menschenrecht, sondern bildet auch das Fundament liberaler Demokratien. Jedoch scheint bei einem größer werdenden Teil der Bevölkerung die Besorgnis um dieses Grundrecht zu wachsen – auch an Universitäten, an denen die freie Rede einen ganz besonderen Stellenwert einnimmt. Allerdings fehlt es trotz intensiver Debatten um „Political Correctness“ und „Cancel Culture“ bislang an Studien, die die Wahrnehmungen von Meinungsfreiheit durch Studierende an deutschen Universitäten sowie ihre Beteiligung an kontroversen Diskussionen auf breiter Basis empirisch untersuchen. Vor diesem Hintergrund geht der vorliegende Beitrag erstmals auf breiterer empirischer Basis der Frage nach, (1) wie Studierende die Meinungsfreiheit an deutschen Universitäten wahrnehmen, (2) wie intensiv sie sich an Diskussionen zu kontroversen Themen beteiligen und welche individuellen Merkmale (3) ihre Wahrnehmungen universitärer Meinungsfreiheit und (4) ihre Redebereitschaft in kontroversen Diskussionen beeinflussen. Die Ergebnisse einer quantitativen Online-Befragung einer disziplinär und regional heterogenen Stichprobe von Studierenden an deutschen Universitäten ($N = 897$), erhoben im Sommer 2023, zeigen, dass Studierende ein hohes Maß an Meinungsfreiheit in ihren Lehrveranstaltungen wahrnehmen. Allerdings finden kontroverse Diskussionen eher selten und vor allem in bestimmten Disziplinen statt und nur eine Minderheit von Studierenden beteiligt sich überhaupt daran. Zudem wird deutlich, dass die Wahrnehmung universitärer Meinungsfreiheit durch politische Einstellungen, Ängste vor Sanktionen durch Mitstudierende und Lehrende, aber auch die politische Kongruenz zu Mitstudierenden geprägt wird. Ob sich Studierende an kontroversen Diskussionen beteiligen, wird dagegen eher von politischem Interesse und Einstellungen sowie befürchteten Sanktionen durch Mitstudierende beeinflusst, wenn auch schwach. Mögliche Implikationen und Anknüpfungspunkte für zukünftige Forschung werden diskutiert.

Schlagwörter: Meinungsfreiheit, Universität, Cancel Culture, Schweigespirale, Redebereitschaft

Abstract: Freedom of speech is not only a fundamental human right but also the foundation of liberal democracies. However, there seems to be growing concern about this basic right among an increasing proportion of the population – including at universities, where free speech is of particular importance. However, despite intense debates about ‘political correctness’ and ‘cancel culture’, there is a lack of studies that empirically examine perceptions of freedom of speech and willingness to speak out in classroom discussions among students at German universities. Against this background, this article examines for the first time on a broader empirical basis (1) how students perceive freedom of expression at German universities, (2) how much they participate in discussions on controversial topics and which individual-level characteristics influence (3) their perceptions of freedom of speech and (4) their willingness to speak out in such discussions. The results of a quantitative online survey of a disciplinary and regionally heterogeneous sample of students at German universities ($N = 897$), conducted in summer 2023, show that students generally perceive a high degree of freedom of expression. However, controversial discussions are rare, the frequency of such debates varies between scientific disciplines, and only few students actually participate. Additionally, perception of academic freedom of speech is shaped by political attitudes, fears of sanctions from fellow students and faculty, as well as political alignment with peers. In contrast, whether students engage in controversial discussions is primarily influenced by political interest and political attitudes as well as anticipated sanctions from fellow students, although these influences are relatively weak. Potential implications and avenues for future research are discussed.

Keywords: Freedom of speech, university, cancel culture, spiral of silence, willingness to speak out

1. Einleitung

Seit einigen Jahren findet in Deutschland eine intensive mediale, gesellschaftliche und wissenschaftliche Debatte um den Stand der Meinungsfreiheit statt. Trotz umfassender verfassungsrechtlicher Garantien und eines hohen Maßes an Meinungs- und Wissenschaftsfreiheit, das Deutschland in internationalen Vergleichen attestiert wird (Freedom House, 2023; Kinzelbach et al., 2024), weisen demoskopische Befunde darauf hin, dass in der Bevölkerung zunehmend Einschränkungen der Meinungsfreiheit wahrgenommen werden (Petersen, 2021b; Rothut et al., 2023; Unzicker, 2018). Eine ähnliche Entwicklung zeigt sich in den USA, deren Ursachen nicht geklärt sind (Gibson & Sutherland, 2023). Die Brisanz dieser Befunde liegt in der Annahme, dass subjektive Wahrnehmungen zunehmend eingeschränkter Meinungsfreiheit mit einer tatsächlichen Zurückhaltung individueller Meinungsäußerungen einhergehen, also Indikatoren für eine geringere politische Redebereitschaft (Matthes et al., 2018; Wilnat & Detenber, 2002) bzw. eine stärkere Selbstzensur sind, die eine Gefährdung des demokratischen Diskurses bedeuten können (Bar-Tal, 2017; Gibson & Sutherland, 2023).

Einen wichtigen Bezugspunkt der Debatte bilden die Universitäten (Breuer, 2019; Eppelsheim, 2022; Erich, 2020; Mansour, 2023). Nach den Anschlägen des 7. Oktober 2023 in Israel, dem anschließenden Krieg in Gaza und den daraus folgenden Studierendenprotesten haben Fragen der Meinungs- und Wissenschaftsfreiheit an deutschen Universitäten noch einmal an Brisanz gewonnen (dazu etwa Eikmanns & Çay, 2024; Grimm et al., 2025; Schieritz & Scholz, 2024). Diese

jüngsten Ereignisse und Debatten stehen jedoch nicht im Fokus des vorliegenden Beitrags, da er auf einer Befragung von Studierenden im Sommer 2023 basiert.

In den bisherigen Diskussionen um die Auswirkungen einer vermeintlichen „Cancel Culture“ an Universitäten, die durch „Political Correctness“, „Wokeness“ bzw. ein linkes Meinungsklima dominiert würden (dazu z. B. Kostner, 2022; Özmen, 2021), spielen Befragungen von Lehrenden an deutschen Hochschulen (Petersen, 2020, 2021a) sowie von Studierenden der Universität Frankfurt (Revers & Traunmüller, 2020) eine wichtige Rolle. Während die empirische Befundlage für die Seite der Lehrenden durch zwei ältere (Petersen, 2020, 2021a) und eine aktuelle Studie (ZEIT-Stiftung, 2024) recht breit ist, lag für die Seite der Studierenden lange keine größer angelegte Studie vor. Dies hat sich erst jüngst mit der Studie von Diehl et al. (2025) geändert, die untersucht, wann deutsche Studierende Einschränkungen akademischer Freiheit zustimmen.

Insbesondere vor dem 7. Oktober 2023 standen sich in der Diskussion um Meinungs- bzw. Wissenschaftsfreiheit an deutschen Universitäten zwei Positionen gegenüber: Die eine kritisierte eine Verengung des Meinungskorridors, die vorrangig von links ausginge. So seien im Namen von Gleichstellung und Antidiskriminierung abweichende Meinungen unterdrückt worden, etwa durch Sprachvorgaben, aber auch durch Proteste, Ausladungen oder Absagen von Veranstaltungen (Berman, 2022; Kostner, 2022). Die andere Seite betrachtete solche Vorfälle eher als medial hochgespielte Einzelereignisse. Nicht Einschränkungen des Sagbaren seien das Problem, sondern dass marginalisierte Gruppen auch an Universitäten nach wie vor zu wenig Gehör fänden und Diskriminierungen ausgesetzt seien (Hoppe et al., 2018; Vogel, 2019).

Die Debatte um Meinungsfreiheit gewinnt im Kontext von Universitäten zusätzliche Brisanz, weil sie als Ort des freien Austauschs von Ideen gelten (z. B. Özmen, 2021). Universitäten stellen einen Lern- und Erfahrungsraum dar, in dem auch kontrovers diskutiert werden muss, nicht zuletzt, um Studierende dazu zu befähigen, fundierte Meinungen zu bilden, diese argumentativ zu vertreten und zugleich Toleranz zu üben (Haake, 2021; Revers & Traunmüller, 2020; Zhou & Barbaro, 2023). Wahrgenommene Einschränkungen der Meinungsfreiheit können dabei nicht nur auf Probleme der akademischen Debattenkultur, sondern des demokratischen Diskurses insgesamt hinweisen (Norris, 2023; Rothut et al., 2023). Dennoch liegt in Deutschland bislang keine breitere empirische Studie dazu vor, wie Studierende die Meinungsfreiheit an Universitäten einschätzen.

Vor diesem Hintergrund verfolgt dieser Beitrag das Ziel, zu erkunden, (1) wie Studierende die Meinungsfreiheit an ihren Universitäten wahrnehmen, (2) inwieweit sie sich im Kontext von Lehrveranstaltungen in kontroversen Debatten zu aktuellen Themen einbringen und (3) welche individuellen Merkmale ihre Wahrnehmungen der universitären Meinungsfreiheit sowie (4) ihre Redebereitschaft beeinflussen. Ein besonderes Augenmerk richten wir einerseits auf die möglichen Auswirkungen politischer Kongruenz bzw. Inkongruenz zwischen Studierenden und Mitstudierenden sowie Lehrenden, andererseits auf mögliche Effekte von Sanktionsängsten, also Befürchtungen, von Mitstudierenden und/oder Lehrenden (sozial) sanktioniert zu werden.

2. Meinungsfreiheit an deutschen Universitäten aus Sicht von Studierenden

2.1 Wahrnehmung von Meinungsfreiheit und die Beteiligung an Diskussionen

Meinungsfreiheit ist ein grundlegendes Menschenrecht und stellt eine zentrale Voraussetzung für das Funktionieren moderner Demokratien dar (Garton Ash, 2016; Quiring et al., 2020). In Deutschland ist das Recht auf freie Meinungsäußerung in Art. 5 Abs. 1 GG verankert (Hillgruber, 2016; Quiring et al., 2020), während Art. 5 Abs. 3 GG die Wissenschaftsfreiheit garantiert (Jaster & Keil, 2021). Im universitären Kontext kann beides relevant sein (Haake, 2021), wobei im Fokus dieser Studie außerwissenschaftliche Meinungsäußerungen stehen, die im Rahmen von Lehrveranstaltungen vorkommen.

Über die rechtliche Dimension hinaus kann Meinungsfreiheit auch aus einer gesellschaftlichen Perspektive betrachtet werden (Sacher & Reinemann, 2025). Sie meint dann die Freiheit, nicht durch soziale Normen, Sanktionen oder Strukturen in der Meinungsäußerung eingeschränkt zu werden (Lanius, 2020; Quiring et al., 2020; Sacher & Reinemann, 2025). Im Folgenden wird Meinungsfreiheit aus dieser Perspektive betrachtet und ihre Wahrnehmung – in Anlehnung an Rothut et al. (2023) – als Realitätsvorstellung konzeptualisiert. Der Fokus liegt somit auf der subjektiven Wahrnehmung sozialer Realität und der individuellen Vorstellung vom „moralisch Sagbaren“ (Lanius, 2020; Rothut et al., 2023, S. 55). Wie Meinungsfreiheit wahrgenommen wird, hängt dabei von individuellen Prädispositionen, persönlichen Erfahrungen sowie medial vermittelten Eindrücken ab (Meltzer, 2017, S. 18; Rothut et al., 2023, S. 55). In diesem Beitrag werden die beiden ersten Faktoren beleuchtet.

Über die diesbezügliche Situation an deutschen Universitäten wissen wir allerdings nur sehr wenig. Aktuelle Studien betrachten das Thema vor allem aus Sicht der Lehrenden (Petersen, 2020, 2021a; ZEIT-Stiftung, 2024). Zwar liegen für andere Länder, insbesondere die USA (Knight Foundation & Ipsos, 2022; Naughton et al., 2017; Sachs, 2018; Stevens & Haidt, 2018a; Zhou & Barbaro, 2023) und Großbritannien (Joint Committee on Human Rights, 2018; YouGov, 2018), umfangreiche Untersuchungen zur Wahrnehmung der universitären Meinungsfreiheit durch Studierende vor. Diese lassen sich aufgrund der Unterschiede zwischen Ländern und Universitätssystemen jedoch nur bedingt auf den deutschen Kontext übertragen (Barendt, 2010; Berman, 2022). Tatsächlich gibt es zur Wahrnehmung der Meinungsfreiheit auf dem Campus durch deutsche Studierende nach unserem Wissen bislang nur eine Studie: Revers und Traummüller (2020) untersuchten den „most likely case“ der politisch mehrheitlich linksorientierten Studierenden an der sozialwissenschaftlichen Fakultät der Universität in Frankfurt. Trotz der von den Autoren von vornherein selbst eingeräumten Einschränkungen (u. a. politische Verzerrung der Stichprobe, geringe Rücklaufquote, Positionalität der Forschenden; vgl. Meier, 2021; Traummüller & Revers, 2021; Villa et al., 2021) ist die Studie ein wichtiger Ansatzpunkt für den vorliegenden Beitrag, da sie sich ebenfalls auf den Kontext von Lehrveranstaltungen konzentriert. So gab die Mehrheit (66 %) der Studierenden an, sich (eher oder völlig) wohlfühlen, die Meinung zu kontroversen politischen Themen in Seminardiskussionen zu äußern.

Allerdings fühlte sich auch ein Drittel (eher oder völlig) gehemmt (35 %) bzw. war (mäßig oder stark) besorgt darüber, von Mitstudierenden oder Dozierenden kritisiert zu werden (31 %), wenn sie sich äußerten. Die Autoren sehen Hinweise dafür, dass an der untersuchten Fakultät ein Konformitätsdruck herrscht, durch den Studierende sich in ihrer Meinungsäußerung gehemmt fühlen. Da bislang aber keine Befunde zur Wahrnehmung der Meinungsfreiheit an deutschen Universitäten über verschiedene Bundesländer, Universitäten und Fächer hinweg vorliegen, formulieren wir folgende erste Forschungsfrage:

FF1: Wie nehmen deutsche Studierende die Meinungsfreiheit an ihren Universitäten wahr?

Bei Revers & Traunmüller (2020) – und unserem Eindruck nach auch in der öffentlichen Debatte – stehen immer wieder bestimmte Fächergruppen im Fokus (vor allem Geistes- und Sozialwissenschaften). Angesichts der zum Teil deutlich unterschiedlichen studentischen Fachkulturen (z. B. Armingeon, 2001; Eßer, 2023), untersuchen wir außerdem, ob es Unterschiede in der Wahrnehmung der Meinungsfreiheit zwischen Studierenden unterschiedlicher Disziplinen gibt. Wir formulieren deshalb als zweite Forschungsfrage:

FF2: Wie unterscheiden sich deutsche Studierende unterschiedlicher Fächergruppen in ihrer Wahrnehmung der Meinungsfreiheit an ihren Universitäten?

Darüber hinaus liegen unserer Kenntnis nach bislang auch keine Befunde dazu vor, wie häufig in den Lehrveranstaltungen an deutschen Universitäten überhaupt über aktuelle kontroverse Themen diskutiert wird und wie intensiv sich Studierende an solchen Debatten beteiligen. Angesichts des oftmals formulierten Anspruchs an Universitäten, Orte der Kontroverse und des freien Worts zu sein, ist dies einigermaßen überraschend. Zudem erscheinen uns Erkenntnisse hierzu wichtig, um Befunde zur Wahrnehmung der Meinungsfreiheit einordnen zu können. Die individuelle Beteiligung an kontroversen Diskussionen wird in der vorliegenden Studie als Redebereitschaft definiert. Wir formulieren deshalb als dritte Forschungsfrage:

FF3: Wie häufig wird in Lehrveranstaltungen an deutschen Universitäten über kontroverse aktuelle Themen diskutiert und wie häufig beteiligen sich Studierende daran?

Die Konzentration auf Lehrveranstaltungen stellt insofern eine Einschränkung dar, als dass Ereignisse und Situationen unberücksichtigt bleiben, die durchaus einen Einfluss auf den Eindruck von studentischer Meinungsfreiheit ausüben könnten (z. B. Diskussionsveranstaltungen, Diskussionen außerhalb von Kursen, Debatten in studentischen Gremien). Allerdings stellen Lehrveranstaltungen für viele Studierende vermutlich den zentralen Teil ihres Studiums dar, sodass sie zu diesen am ehesten Auskunft geben können. Außerdem geht es uns um einen Eindruck von der alltäglichen akademischen Diskurskultur, weniger um spektakuläre Ereignisse. Wir erwarten dabei auch Unterschiede zwischen Disziplinen, weshalb wir eine weitere Forschungsfrage formulieren:

FF4: Wie unterscheiden sich verschiedene Fächergruppen in der Häufigkeit der Diskussion kontroverser aktueller Themen und der Redebereitschaft der Studierenden?

2.2 Einflüsse von politischer Kongruenz auf die Wahrnehmung von Meinungsfreiheit und die Redebereitschaft

Studien, die die Wahrnehmung der Meinungsfreiheit im universitären Kontext erklären wollen, beziehen sich in der Regel auf die Theorie der öffentlichen Meinung (Matthes et al., 2018; Noelle-Neumann, 1974, 1980). Demnach beobachten Menschen kontinuierlich ihr soziales Umfeld, da ihnen das Grundbedürfnis inhärent ist, zu einer Gruppe zu gehören und soziale Beziehungen aufrechtzuerhalten (Baumeister & Leary, 1995). Diese Beobachtung ermöglicht es, Schwankungen der öffentlichen Meinung und die der Mehrheit zu ermitteln. Die Theorie geht weiter davon aus, dass Menschen dazu neigen, ihre Meinung eher öffentlich zu äußern, wenn sie diese als kongruent zur von ihnen wahrgenommenen Mehrheitsmeinung empfinden, und weniger dazu neigen, sich zu äußern, wenn sie ihre Meinung als inkongruent zur wahrgenommenen Mehrheitsmeinung empfinden. Das Schweigen aus Angst vor Isolation oder sozialen Sanktionen wird auch als Selbstzensur bezeichnet und kann dazu führen, dass die scheinbare Mehrheitsmeinung gestärkt und die Minderheit „in der öffentlichen Meinung im Laufe der Zeit immer weniger sichtbar“ (Neubaum & Krämer, 2018, S. 139) wird, was als Prozess der Schweigespirale bezeichnet wird (Neubaum & Krämer, 2018; Noelle-Neumann, 1974).

Im universitären Kontext resümieren Revers und Traunmüller (2020) bei ihrer Betrachtung eines eher linken politischen Umfelds, dass politisch weiter rechtsstehende Studierende gehemmter waren, ihre Meinungen zu kontroversen Themen zu äußern. Als Erklärung führen die Autoren mit Verweis auf die Theorie der Schweigespirale an, dass sich rechtsgerichtete Studierende in dem eher linksorientierten Klima auf dem Campus in der Minderheit fühlen und deshalb eine negativere Meinungsfreiheitswahrnehmung und auch eine geringere Redebereitschaft aufweisen könnten (Revers & Traunmüller, 2020, S. 491). Dies legt nahe, dass die individuelle Übereinstimmung mit der wahrgenommenen Orientierung Mitstudierender, die als politische Kongruenz bezeichnet werden kann, ein entscheidender Faktor für die Wahrnehmung von Meinungsfreiheit und die Redebereitschaft ist (Noelle-Neumann, 1974, 1980; Norris, 2023).

Aufgegriffen wurden die Theorie der öffentlichen Meinung und das Konzept der politischen Kongruenz mit Blick auf den universitären Kontext auch von Norris (2023). Sie untersuchte mit Daten einer international vergleichenden Befragung unter anderem, ob Politikwissenschaftler und Politikwissenschaftlerinnen eine Veränderung des Meinungsklimas hin zu mehr „Political Correctness“ und eine sich ausbreitende „Cancel Culture“ wahrnahmen. Norris (2023) zeigt, dass Befragte in liberalen Demokratien (z. B. den USA oder Deutschland), die sich weiter rechts im politischen Spektrum verorteten, eine zunehmende „Cancel Culture“ wahrnahmen. In eher konservativ geprägten Ländern (z. B. Nigeria) war diese

Wahrnehmung dagegen stärker unter Befragten verbreitet, die sich als eher links und progressiv betrachteten. Norris (2023) schließt im Einklang mit der Theorie der Schweigespirale, dass Diskrepanzen zwischen individuellen Einstellungen und dem dominierenden Meinungsklima nicht nur zur Wahrnehmung von Meinungsdruck führen, sondern auch die Gefahr von Selbstzensur erhöhen. Allerdings wurde die Wahrnehmung des Meinungsklimas hier nicht auf Individualebene gemessen.

Anders ist dies in der Studie von Traummüller (2022, 2023), der Daten einer großen Studierendenbefragung aus den USA analysierte. Abhängige Variable ist hier zwar nicht die Wahrnehmung von Meinungsfreiheit, sondern die Neigung zur Selbstzensur. Die Befunde sind hier dennoch relevant, weil sie zeigen, dass die Neigung zur Selbstzensur bei Studierenden einerseits von ihrer eigenen politischen Orientierung und andererseits von der auf dem Campus wahrgenommenen vorherrschenden politischen Orientierung abhängt: So neigen konservative Studierende in einem liberal dominierten Universitätskontext eher und in einem konservativen Kontext weniger dazu, ihre eigene Meinung aufgrund befürchteter Sanktionen zurückzuhalten (Traummüller, 2022, 2023).

Kaum berücksichtigt wird bislang die Tatsache, dass Studierende an ihren Universitäten mit sehr unterschiedlichen Bezugsgruppen konfrontiert sind, die jeweils eigene Quellen für Wahrnehmungen politischer Kongruenz oder Inkongruenz darstellen können. Zu nennen sind hier vor allem Mitstudierende und Lehrende. Mitstudierende können als Teil der eigenen Peer-Group betrachtet werden, die als soziales Umfeld für die Studierenden von erheblicher Bedeutung ist. Vor dem Hintergrund der Forschung zur Schweigespirale, die nahelegt, dass in Familie und Freundeskreis die stärksten Schweigeeffekte auftreten (Matthes et al., 2018, S. 23; dazu auch Cook & Heilmann, 2013), würde man erwarten, dass von dieser sozialen Gruppe ein besonders starker normativer Einfluss ausgeht. Tatsächlich ergab eine Studie aus den USA, dass die Zurückhaltung von Studierenden, sensible Themen im universitären Kontext zu diskutieren, in erster Linie auf Ängste vor den möglichen Reaktionen von Mitstudierenden zurückzuführen war (Zhou & Barbo, 2023).

Zu Lehrenden stehen Studierende dagegen in einem asymmetrischen Machtverhältnis (Cialdini & Goldstein, 2004). Lehrende haben zumindest potenziell die Möglichkeit, ihre Machtposition zum Nachteil der Studierenden auszunutzen und missliebige Meinungsäußerungen (z. B. durch Bloßstellung oder gar eine negative Bewertung von Prüfungsleistungen) zu sanktionieren. Allein schon die Vermutung, dass dies möglich sein könnte, mag den Eindruck eingeschränkter Meinungsfreiheit befördern oder Studierende dazu verleiten, ihre Meinungen zurückzuhalten. Erleben Studierende sogar, dass Lehrende auf Meinungsäußerungen negativ oder mit Abwertung reagieren, dürfte dieser Effekt umso größer ausfallen. Entsprechende Negativbeispiele finden sich etwa bei Revers und Traummüller (2020). Vor diesem Hintergrund formulieren wir unsere ersten Hypothesen:

H1: Je höher die politische Kongruenz zwischen Studierenden und (a) ihren Mitstudierenden sowie (b) ihren Lehrenden, desto positiver wird die universitäre Meinungsfreiheit wahrgenommen.

H2: Je höher die politische Kongruenz zwischen den Studierenden und (a) ihren Mitstudierenden sowie (b) ihren Lehrenden, desto eher beteiligen sie sich an universitären Diskussionen über kontroverse Themen.

2.3 Einflüsse von Sanktionsängsten auf die Wahrnehmung von Meinungsfreiheit und die Beteiligung an Diskussionen

Ein weiterer Faktor, der sich im Zusammenhang mit der allgemeinen Wahrnehmung von Meinungsfreiheit als relevantes Konstrukt erwiesen hat, sind erwartete Reaktionen bzw. Sanktionen, die als Folgen einer Meinungsäußerung befürchtet werden. Sie spielen in der Theorie der öffentlichen Meinung (Noelle-Neumann, 1974, 1980), aber auch in der Forschung zu sozialem Ausschluss (Williams, 2007) eine vermittelnde Rolle zwischen der Wahrnehmung des Meinungsklimas und der eigenen Meinungsäußerung. Die Erwartung von Reaktionen bzw. Sanktionen geht dabei mit einer stärkeren wahrgenommenen Einschränkung der Meinungsfreiheit (Rothut et al., 2023) und einer geringeren Redebereitschaft einher (Matthes et al., 2018; Neubaum & Krämer, 2018). Die in der Forschung diskutierten Reaktionen bzw. Sanktionen umfassen ein breites Spektrum, das von der Erwartung, sich rechtfertigen zu müssen, über die Befürchtung von Widerspruch und Bloßstellung bis hin zum Ausschluss aus einer Gruppe und beruflichen Nachteilen reichen kann (Neubaum & Krämer, 2018). Auch Revers und Traunmüller (2020) stellen fest, dass sich konservativ orientierte Studierende eher Sorgen über die Folgen ihrer politischen Meinungsäußerungen machen, also eine größere Sanktionsangst aufwiesen.

Man kann deshalb annehmen, dass Studierende, die befürchten, von ihren Mitstudierenden oder ihren Lehrenden negative Reaktionen zu erhalten und damit (sozial) sanktioniert zu werden, die universitäre Meinungsfreiheit eher als eingeschränkt wahrnehmen und weniger dazu neigen, sich in Diskussionen zu äußern. Allerdings dürfte sich das Spektrum erwarteter negativer Reaktionen bzw. die Möglichkeit zur Sanktionierung von Meinungsäußerungen für Mitstudierende und Lehrende durchaus unterscheiden, da hier eine Machtasymmetrie vorliegt. Wir formulieren deshalb zwei weitere Hypothesen:

H3: Je eher Studierende erwarten, bei Meinungsäußerungen (a) von ihren Mitstudierenden und (b) von ihren Lehrenden negative Reaktionen bzw. Sanktionen zu erfahren, desto negativer werden sie die universitäre Meinungsfreiheit wahrnehmen.

H4: Je eher Studierende erwarten, bei Meinungsäußerungen (a) von ihren Mitstudierenden und (b) von ihren Lehrenden negative Reaktionen bzw. Sanktionen zu erfahren, desto weniger beteiligen sie sich an universitären Diskussionen über kontroverse Themen.

2.4 Weitere Einflussfaktoren auf die Wahrnehmung von Meinungsfreiheit und die Beteiligung an Diskussionen

Bisherige Studien zur Wahrnehmung von Meinungsfreiheit und zur Redebereitschaft außerhalb des universitären Kontexts legen nahe, dass es weitere individuelle Merkmale gibt, die diese beeinflussen.

Politische Prädispositionen. Mehrere Studien sprechen für eine hohe Bedeutung politischer Prädispositionen. So zeigen Rothut et al. (2023), dass politische Selbsteinstufung, Parteipräferenz und Politikverdrossenheit einen starken Einfluss auf die generelle Wahrnehmung der Meinungsfreiheit haben. In ihrer Studie nehmen politikverdrossene Personen, Personen auf der rechten Seite des politischen Spektrums sowie Anhängerinnen und Anhänger der AfD die Meinungsfreiheit als besonders eingeschränkt wahr (ähnlich Menzner & Traummüller, 2023). Fraglich ist, ob diese Faktoren auch für die studentische Wahrnehmung universitärer Meinungsfreiheit eine Rolle spielen. Zum einen dürften Studierende als formal hochgebildete und eher privilegierte Gruppe eine relativ geringe Politikverdrossenheit aufweisen (Shell Jugendstudie, 2019). Allerdings ist denkbar, dass Studierende mit höherer Politikverdrossenheit deshalb eine geringere politische Kongruenz zu Mitstudierenden wahrnehmen (Heil, 2016) und sich dadurch eher abseits vom „Mainstream“ verorten. Quiring et al. (2020) zeigen darüber hinaus, dass geringere Demokratiezufriedenheit zu einer stärkeren Wahrnehmung von einschränkenden Sprechverböten führte. Wir werden die möglichen Einflüsse der genannten Prädispositionen deshalb kontrollieren.

Im Hinblick auf die Redebereitschaft erscheint nicht ausgeschlossen, dass sich Personen mit extremen politischen Positionen auch dann äußern, wenn sie sich in einer Minderheitenposition sehen. Entsprechende Hinweise finden sich nicht nur in der Theorie der öffentlichen Meinung, wo entsprechende Gruppen als „harter Kern“ oder „Avantgarde“ beschrieben werden (Noelle-Neumann, 1974, 1980), sondern auch in der Forschung zum Zusammenhang politischer Einstellungen und Redebereitschaft (Burnett et al., 2022) oder zu Meinungsäußerungen in sozialen Medien (Wintterlin et al., 2021).

Sonstige Merkmale: Einschlägige Studien zeigen weiterhin, dass es persönlichkeitspsychologische Dispositionen gibt, die einen Einfluss auf die Wahrnehmung der Meinungsfreiheit (z. B. Rothut et al., 2023; Menzner & Traummüller, 2023) bzw. die Redebereitschaft ausüben (können) (Matthes et al., 2018; Willnat et al., 2002). Wir kontrollieren deshalb die Merkmale Alter, Geschlecht und politisches Interesse.

3. Methodisches Vorgehen

3.1 Design und Stichprobe

Um die Forschungsfragen zu beantworten und die Hypothesen zu testen, wurde eine standardisierte quantitative Online-Befragung durchgeführt. Befragt wurden Studierende an deutschen Universitäten, die im Online-Access-Panel des kommerziellen Dienstleisters Bilendi & Respondi registriert waren. Es handelt sich also

nicht um eine Zufallsstichprobe, deren Ziehung in Ermangelung eines Studierendenregisters äußerst aufwendig wäre. Daher nehmen wir befragungsbereite Studierende in den Blick und stellen sicher, dass diese den für unsere Fragestellungen wichtigen Merkmalen der Grundgesamtheit der Studierenden entsprechen (siehe unten). Zudem hat eine aktuelle Methodenstudie gezeigt, dass Verteilungen wichtiger Eigenschaften von Studierenden in Online-Access-Panels (z. T. nach entsprechender Gewichtung) sehr ähnlich zu denen in wahrscheinlichkeitsbasierten Studierendenbefragungen sind (Hinz et al., 2023). Obwohl wir deshalb davon ausgehen, dass unser Vorgehen eine gute Annäherung an die Wahrnehmungen und Verhaltensweisen deutscher Studierender sicherstellt, sollten die berichteten Prozentwerte vorsichtig interpretiert werden. Die späteren Zusammenhangsanalysen (Regressionen) sollten dagegen vergleichsweise robust gegenüber Verzerrungen sein. Dennoch haben wir zahlreiche Zusatzanalysen durchgeführt, um diese Robustheit auch empirisch zu prüfen (siehe unten). Der Fragebogen wurde mit dem Online-Befragungs-Tool SoSci Survey erstellt und drei Pretests unterzogen. Die Feldphase erstreckte sich vom 17.7.2023 bis zum 2.8.2023. Um eine hohe Teilnahmequote sicherzustellen, wurden im Fragebogen zunächst keine weiteren Ausschlusskriterien gewählt oder Quotierungen vorgenommen.

Insgesamt füllten 1.098 Studierende aus 16 Bundesländern und von über 99 verschiedenen Universitäten den Fragebogen vollständig aus. Um eine hohe Qualität der Daten sicherzustellen, wurde in den Fragebogen erstens ein Item zur Prüfung der Aufmerksamkeit integriert. Zudem wurden anhand des Relative-Speed-Index Personen ausgeschlossen, die den Fragebogen zu schnell beantwortet hatten (Leiner, 2019). Personen, die nicht in Deutschland studierten, Promotionsstudierende sowie Fälle mit unplausiblen Antworten wurden aus dem Datensatz entfernt. Zudem wurden aufgrund der geringen Gruppengröße neun Studierende von der Analyse ausgeschlossen, die als Geschlecht divers angegeben hatten. Insgesamt wurden somit 180 Fragebögen entfernt.

Da es sich nicht um eine Zufallsauswahl handelt, haben wir die Stichprobe anhand der Merkmale Geschlecht und Studiengang mit der Grundgesamtheit der Studierenden an deutschen Universitäten verglichen. Hierfür wurden zunächst die ursprünglich erhobenen Einzelfächer unter Rückbezug auf die Fächersystematik des Statistischen Bundesamtes (Statistisches Bundesamt, 2024) sowie basierend auf eigenen Zuordnungen zu sechs Fächergruppen zusammengefasst. Ein detaillierter Überblick dazu findet sich in Anhang A auf OSF. Basierend auf der Studierendenstatistik aus dem Wintersemester 2022/2023 (Statistisches Bundesamt, 2023) wurde die ungewichtete Stichprobe unter Berücksichtigung der Geschlechterverteilung innerhalb der Fächergruppen mit der Grundgesamtheit verglichen. Hierbei zeigten sich zum einen Abweichungen im Hinblick auf die Stärke einzelner Fächer (Anhang B auf OSF: Tabelle B.1, Tabelle B.2). In der ursprünglichen Stichprobe waren insbesondere Studierende der Ingenieurwissenschaften deutlich unterrepräsentiert und Studierende der Geisteswissenschaften & Kunst stark überrepräsentiert. Zudem war der Anteil weiblicher Studierender deutlich höher als in der Grundgesamtheit. Um den Verzerrungen entgegenzuwirken, wurde eine nachträgliche Gewichtung (Poststratifizierung) vorgenommen (Sand & Kunz, 2020). Neuere Studien zeigen, dass eine solche Gewichtung von Daten aus On-

line-Access-Panels im Hinblick auf interessierende Merkmale eine gute Annäherung an die Werte von Zufallsstichproben ermöglichen kann (Grewenig et al., 2023). Für die Gewichtung wurden sowohl die Gesamtverteilung der Fächergruppen als auch die Geschlechterverteilung innerhalb der einzelnen Fächergruppen berücksichtigt (siehe Tabelle B.2 für Details und Gewichtungsfaktoren). Insbesondere bei der Interpretation von Anteilswerten bleibt zu berücksichtigen, dass eine solche Gewichtung die ursprüngliche Verzerrung zwar reduzieren, jedoch nicht vollständig ausgleichen kann (Sand & Kunz, 2020).

Die endgültige gewichtete Stichprobe umfasst rechnerisch 897 Studierende aus sechs Fächergruppen ($N = 897$). Davon sind 52 % weiblich und 48 % männlich. Im Durchschnitt sind die Studierenden 25 Jahre alt ($SD = 4.93$). Von den Befragten streben 50 % einen Bachelorabschluss an, 28 % einen Masterabschluss und 20 % ein Staatsexamen. Den größten Anteil in der endgültigen Stichprobe stellen die Geisteswissenschaften & Kunst dar (20,1 %), gefolgt von Rechts- & Wirtschaftswissenschaften (18,9 %), Ingenieurwissenschaften (17,7 %), Mathematik, Natur- & Agrarwissenschaften (17,3 %), Sozialwissenschaften (15,4 %) sowie Medizin, Gesundheits- & Sportwissenschaften (10,5 %). Auch die Verteilung auf die Bundesländer entspricht, abgesehen von kleineren Abweichungen, der Struktur der Grundgesamtheit. Alle folgenden Angaben basieren auf der beschriebenen Gewichtung.

3.2 Operationalisierung

3.2.1 Abhängige Variablen

Wahrgenommene Meinungsfreiheit: Um zu erfassen, wie die Studierenden den Stand der universitären Meinungsfreiheit im eigenen Hauptfach wahrnehmen, bewerteten sie acht Items auf einer siebenstufigen Skala (siehe Anhang C.1 auf OSF; Beispielitem: „Es gibt viele ungeschriebene Gesetze darüber, welche Meinungen in meinem Hauptfach sozial akzeptabel und zulässig sind und welche tabu“; 1 = stimme überhaupt nicht zu, 7 = stimme voll und ganz zu). Die Items wurden in Anlehnung an Rothut et al. (2023) formuliert, die diese wiederum in Teilen selbst erstellt und in Teilen an Köcher (2019) angelehnt haben. Sie decken unausgesprochene Diskursnormen und Tabus, die Sagbarkeit und die möglichen Folgen von Meinungsäußerungen sowie „Political Correctness“ und Gleichberechtigung ab. Nach dem Umcodieren eines invers formulierten Items wurde für die multivariaten Analysen ein Mittelwertindex ($\alpha = .834$, $M = 2.91$, $SD = 1.21$, $n = 889$) gebildet. Höhere Werte verweisen auf die Wahrnehmung stärkerer Einschränkungen der universitären Meinungsfreiheit. Im Mittel ist diese Wahrnehmung bei den Befragten eher gering ausgeprägt, wobei 68 % keine bis eher geringe und 14 % eher starke bis starke Einschränkungen sehen.

Häufigkeit kontroverser Diskussionen & Redebereitschaft: Um die Beteiligung an kontroversen Diskussionen zu ermitteln, wurden die Studierenden zunächst mittels einer vierstufigen Skala gefragt, wie häufig es in Seminaren oder Vorlesungen ihres Hauptfachs überhaupt zu solchen Diskussionen kommt (siehe auch Anhang C.2 auf OSF; 1 = noch nie; 2 = selten; 3 = gelegentlich; 4 = oft). Dies war im

Mittel „gelegentlich“ der Fall ($M = 2.27$, $SD = .94$, $n = 875$), wobei 23 % der Studierenden angaben, dass in ihrem Hauptfach nie kontrovers diskutiert wird, während 11,5 % angaben, dies sei oft der Fall. Studierende, die angaben, dass in Lehrveranstaltungen ihres Hauptfachs „nie“ kontrovers diskutiert wird, wurden von der weiteren Analyse der Prädiktoren der Redebereitschaft ausgeschlossen, sodass hierfür noch $N = 670$ Befragte zur Verfügung standen.

Die individuelle Beteiligung an kontroversen Diskussionen (Redebereitschaft) wurde dann auf einer fünfstufigen Skala (1 = noch nie, 5 = immer) erhoben. Hier sollten die Studierenden, in deren Lehrveranstaltungen tatsächlich kontroverse Diskussionen stattfanden, angeben, wie häufig sie sich bislang an solchen Diskussionen beteiligt haben ($M = 2.55$, $SD = 1.00$, $n = 670$). Dies war im Durchschnitt „selten“ bis „gelegentlich“ der Fall, wobei sich 48 % „nie“ oder „selten“, 34 % „gelegentlich“ und 18 % „oft“ oder „immer“ beteiligt haben.

3.2.2 Unabhängige Variablen

Politikverdrossenheit: Um die Politikverdrossenheit zu erheben, wurde analog zu Rothut et al. (2023) ein Item zur allgemeinen Demokratiezufriedenheit mit vier Items der Political-Efficacy-Kurzskala von Beierlein et al. (2014) kombiniert. Die Zustimmung wurde auf einer siebenstufigen Skala (1 = „stimme überhaupt nicht zu“, 7 = „stimme voll und ganz zu“) abgefragt. Für eine vereinfachte Interpretation wurden alle Items invertiert. Trotz der eher schwachen Reliabilität ($\alpha = .658$) wurden sie dann zu einem Mittelwertindex verrechnet ($M = 3.49$, $SD = .99$, $n = 894$), da es sich bei Politikverdrossenheit um ein etabliertes Konstrukt handelt (vgl. Rothut et al., 2023). Hohe Werte weisen auf eine stark ausgeprägte Politikverdrossenheit hin, die in der Stichprobe durchschnittlich in mittlerem Maße ausgeprägt ist. Bei 49 % zeigt sich eine eher schwache bis sehr schwache, bei 15 % eine eher hohe bis sehr hohe Politikverdrossenheit.

Demokratiezufriedenheit: Unter Rückgriff auf die Erhebung der Demokratiezufriedenheit bei Decker et al. (2019) beantworteten die Befragten auf einer vierstufigen Skala (1 = überhaupt nicht zufrieden, 4 = sehr zufrieden), inwieweit sie „alles in allem“ mit dem Funktionieren der Demokratie in Deutschland zufrieden sind ($M = 2.78$, $SD = .68$, $n = 867$). Im Mittel zeigen sie sich ziemlich zufrieden, wobei 27 % überhaupt nicht oder wenig zufrieden, 73 % ziemlich oder sehr zufrieden sind.

Politische Orientierung und wahrgenommene politische Orientierung Mitstudierender und Lehrender: Die Abfrage der eigenen politischen Orientierung und der vermuteten politischen Orientierung Mitstudierender sowie Lehrender erfolgte auf einer elfstufigen Links-Rechts-Skala (1 = links, 11 = rechts; GESIS, o. D.). Ihre eigene politische Orientierung schätzen die Studierenden im Durchschnitt etwas links der Mitte ein ($M = 5.11$, $SD = 1.88$, $n = 870$) und damit auch etwas weiter links, als sich die Gesamtbevölkerung (z. B. Infratest dimap, 2024) sowie die Gesamtgruppe der Jungwähler einschätzt (Döbele et al., 2023). 57 % verorten sich dabei links und 19 % rechts der politischen Mitte. Die Fächergruppen unterscheiden sich in ihrer politischen Einstellung zum Teil signifikant voneinander. Am weitesten links schätzen sich Studierende der Sozialwissenschaften ($M = 4.42$)

sowie der Geisteswissenschaften & Kunst ($M = 4.49$) ein, signifikant weiter rechts sehen sich Studierende der Mathematik, Natur- & Agrarwissenschaften ($M = 5.21$), der Medizin, Gesundheits- & Sportwissenschaften ($M = 5.25$) sowie der Rechts- & Wirtschaftswissenschaften ($M = 5.36$). Am ehesten in der Mitte der Skala und damit am weitesten rechts stehen in unserer Stichprobe die Studierenden der Ingenieurwissenschaften ($M = 5.98$). Diese Befunde entsprechen weitgehend bisherigen Erkenntnissen zu den Unterschieden politischer Einstellungen der Studierenden verschiedener Fächer, was auch als weiterer Beleg für die Qualität der Stichprobe gedeutet werden kann (Armingeon, 2001; Lindov, 2020). Die politische Orientierung der Mitstudierenden des eigenen Hauptfachs an der eigenen Universität verorten die Studierenden im Mittel noch etwas weiter links ($M = 4.73$, $SD = 1.83$, $n = 833$), wobei 65 % diese links und 13 % rechts der politischen Mitte einschätzen. Die politische Orientierung der Lehrenden vermuten die Studierenden durchschnittlich etwas weiter rechts als die ihrer Mitstudierenden ($M = 5.42$, $SD = 1.58$, $n = 787$). 46 % verorten sie links der Mitte und 18 % rechts der Mitte.

Parteipräferenz: Die Parteipräferenz der Befragten wurde anhand der Sonntagsfrage gemessen. Der größte Anteil der Studierenden wollte zum Zeitpunkt der Befragung Bündnis 90/Die Grünen wählen (29 %), gefolgt von SPD (14 %), CDU/CSU (11 %), Linke (8 %), FDP (7 %) und AfD (4 %). 5 % wollten für eine sonstige Partei stimmen, 4 % nicht wählen und 18 % nutzten die „Weiß nicht“-Option. Diese Ergebnisse stehen im Einklang mit Befunden der Wahlforschung, wonach beispielsweise Jüngere und höher Gebildete deutlich stärker zu Bündnis 90/Die Grünen neigen als der Durchschnitt der Bevölkerung (Hübner et al., 2022; John, 2021).

Politische Kongruenz mit Mitstudierenden und Lehrenden: Um die Kongruenz ihrer eigenen mit der politischen Orientierung ihrer Mitstudierenden bzw. Lehrenden abzubilden, wurde jeweils der Differenzbetrag aus beiden Einschätzungen gebildet. Zuvor wurden fehlende Werte für die Einschätzung von Mitstudierenden und Lehrenden durch den jeweiligen Mittelwert ersetzt (Imputation), wobei die Auswirkungen der Imputation (und der Gewichtung) auf die Substanz der Befunde in späteren Robustness-Checks überprüft wurden (s. u.). Die Werte wurden anschließend für eine vereinfachte Interpretation invertiert. Die berechnete Variable weist einen Wertebereich von 0 (vollständige politische Inkongruenz) bis 10 (vollständige politische Kongruenz) auf. Betrachtet man die politische Kongruenz mit Mitstudierenden, so fällt diese recht hoch aus: Viele Studierende nehmen ihre eigene politische Orientierung (sehr) ähnlich zu der ihrer Mitstudierenden wahr ($M = 8.43$, $SD = 1.41$, $n = 897$). Die große Mehrheit der Studierenden (95 %) ist über dem Skalenmittelpunkt zu verorten und zeichnet sich damit durch eine (eher) hohe politische Kongruenz aus. 23 % nehmen ihre Einstellung sogar als identisch mit der ihrer Mitstudierenden wahr und niemand nimmt eine völlige Inkongruenz wahr (MIN = 2, MAX = 10). Die Kongruenz mit den eigenen Lehrenden ist im Mittel sogar geringfügig höher ($M = 8.54$, $SD = 1.26$, $n = 897$). Auch hier nimmt keiner der Befragten eine vollständige Inkongruenz wahr (MIN = 2, MAX = 10). 23 % gehen davon aus, dass sich ihre eigene politische

Einstellung mit der ihrer Lehrenden deckt. Während nur 1 % (eher) eine politische Inkongruenz vermuten, weisen 97 % (eher) eine politische Kongruenz auf.

Angst vor Sanktionen durch Mitstudierende und Lehrende: Zur Erfassung der Angst vor Reaktionen bzw. (sozialen) Sanktionen wurden dem Vorgehen von Roth et al. (2023, S. 63) folgend die Dimensionen der Skala von Neubaum und Krämer (2018) aufgegriffen und auf Basis der robusten Ergebnisse reduziert. Anschließend wurden sie inhaltlich auf Studierende und Lehrende übertragen. Die Befragten sollten auf einer Ja-/Nein-Skala angeben, ob sie die beschriebenen Reaktionen bzw. Sanktionen bereits befürchtet und sich deshalb nicht getraut haben, ihre Meinung in einer Vorlesung oder einem Seminar des eigenen Hauptfachs zu äußern. Befürchtungen vor Sanktionen durch Mitstudierende wurden mittels fünf Items erhoben (z. B. „von anderen Studierenden ausgelacht zu werden“; „dass mir ein anderer Studierender widerspricht“, „von anderen Studierenden beleidigt zu werden“), die Angst vor Sanktionen durch Lehrende mittels sechs Items (z. B. „dass mir ein Dozierender widerspricht“, „von Dozierenden ausgelacht zu werden“, „vom Dozierenden verbal angegriffen zu werden“). Die Items wurden dummy-codiert und zu einem Summenindex für die Angst vor Sanktionen durch Mitstudierende bzw. Lehrende addiert, der Werte zwischen 0 (keine Angst vor Sanktionen) und 5 bzw. 6 (starke Angst vor Sanktionen) aufweist. Die Angst vor Sanktionen durch Mitstudierende fällt im Mittel gering aus ($\alpha = .747$, $M = 1.28$, $SD = 1.41$, $n = 897$), wobei 81 % der Studierenden keine bis eher geringe und 10 % eher starke bis sehr starke Angst vor Sanktionen berichten. Vor Sanktionen durch Lehrende haben die Befragten im Durchschnitt minimal mehr, insgesamt aber ebenfalls wenig Angst ($\alpha = .773$, $M = 1.34$, $SD = 1.54$, $n = 897$). Bei 83 % ist die Sanktionsangst (eher) gering ausgeprägt, bei 7 % (eher) stark.

Kontrollvariablen: Als Kontrollvariablen wurden zudem Geschlecht (dichotomisiert; $n = 432$ männlich, $n = 464$ weiblich), Alter ($M = 25.10$, $SD = 4.93$) und politisches Interesse ($M = 4.57$, $SD = 1.64$) einbezogen. Das politische Interesse wurde nach Otto und Bacherle (2011) durch fünf Items (Beispielitem: „Ganz allgemein interessiere ich mich sehr für Politik“) auf einer siebenstufigen Skala (1 = „trifft überhaupt nicht zu“, 7 = „trifft voll und ganz zu“) erfasst, die einen Mittelwertindex bildet ($\alpha = .943$).

3.3 Analysestrategie und Robustness-Checks für Gewichtung und Imputation

Die Hypothesen wurden mit linearen multiplen Regressionen getestet. Abhängige Variablen waren der Index der Wahrnehmung der universitären Meinungsfreiheit bzw. das Einzelitem zur Redebereitschaft bei kontroversen Themen in Lehrveranstaltungen. Da unsere konzeptionell wichtigen unabhängigen Variablen zur politischen Kongruenz und zu Sanktionserwartungen eine Reihe fehlender Werte enthielten, wurden diese Werte näher analysiert. Zunächst zeigte Littles MCAR Test, dass die fehlenden Werte nicht „Missing Completely at Random“ (MCAR) waren. Da ein statistischer Test zur Unterscheidung von „Missings at Random (MAR)“ und „Missing Not at Random (MNAR)“ nicht existiert, wurde die Entscheidung getroffen, die Analysen umfangreich abzusichern (Schafer & Graham, 2002; van Buuren, 2018). Da bei „Missings at Random“ die Beziehungen zwi-

schen den Daten erhalten bleiben, entschieden wir uns zunächst für zwei Varianten von Imputation. Die fehlenden Werte wurden dabei einmal mit dem Mittelwert der Gesamtstichprobe und einmal mittels multipler Imputation (FCS; linear) ersetzt. In der multiplen Imputation wurden zur Vorhersage fehlender Werte, den Empfehlungen der Literatur folgend, alle Variablen einbezogen, die auch in den späteren Regressionsanalysen verwendet wurden (Schafer & Graham, 2002; van Buuren, 2018). Um auch der Möglichkeit Rechnung zu tragen, dass es sich tatsächlich um „Missings Not at Random“ handelt und eine Imputation zu Verzerrungen der Befunde führt, wurden zudem alle Analysen einmal mit und einmal ohne Imputationen der fehlenden Werte durchgeführt. Dieses Vorgehen trägt der inhärenten Unsicherheit im Hinblick auf die Ursachen der fehlenden Werte umfassend und transparent Rechnung.

Um die Robustheit der Befunde zu prüfen, wurden die Regressionsmodelle dann einmal mit und einmal ohne Gewichtung sowie mit bzw. ohne die verschiedenen Imputationen berechnet (Robustness-Checks; siehe Anhang D auf OSF). Für die beiden abhängigen Variablen wurden zunächst jeweils ein Modell mit den Originaldaten (ohne Gewichtung und ohne Imputation; Modell 1 bzw. A), ein Modell mit Gewichtung und ohne Imputation (Modell 2 bzw. B), ein Modell ohne Gewichtung und mit Mittelwert-Imputation (Modell 3 bzw. C) sowie ein Modell mit Gewichtung und Mittelwert-Imputation berechnet (Modell 4 bzw. D) (siehe Anhänge D.1 und D.2). In einem zweiten Schritt wurde dann eine multiple Imputation durchgeführt und die Analysen für beide Variablen mit und ohne Gewichtung wiederholt. Da im Rahmen der multiplen Imputation jeweils zehn Alternativmodelle erstellt wurden, ergeben sich für jede abhängige Variable jeweils 20 weitere Regressionsmodelle, die zur Beurteilung der Robustheit der Befunde herangezogen werden (Modelle Mult 1–20 bzw. Mult A–T; siehe Anhänge D.3–D.6). Somit wurden für jede abhängige Variable 24 Regressionsmodelle berechnet. Deren Befunde sind für die Wahrnehmung der Meinungsfreiheit außerordentlich robust, für die Redebereitschaft nur etwas weniger robust. Für eine bessere Übersichtlichkeit bezieht sich die folgende Ergebnisdarstellung auf die Modelle mit gewichteten Daten und Mittelwert-Imputation (Modelle 4 bzw. D), wobei die Robustheit der Einflüsse über alle Modelle hinweg jeweils thematisiert wird.

4. Ergebnisse

4.1 Wahrnehmung der Meinungsfreiheit an deutschen Universitäten

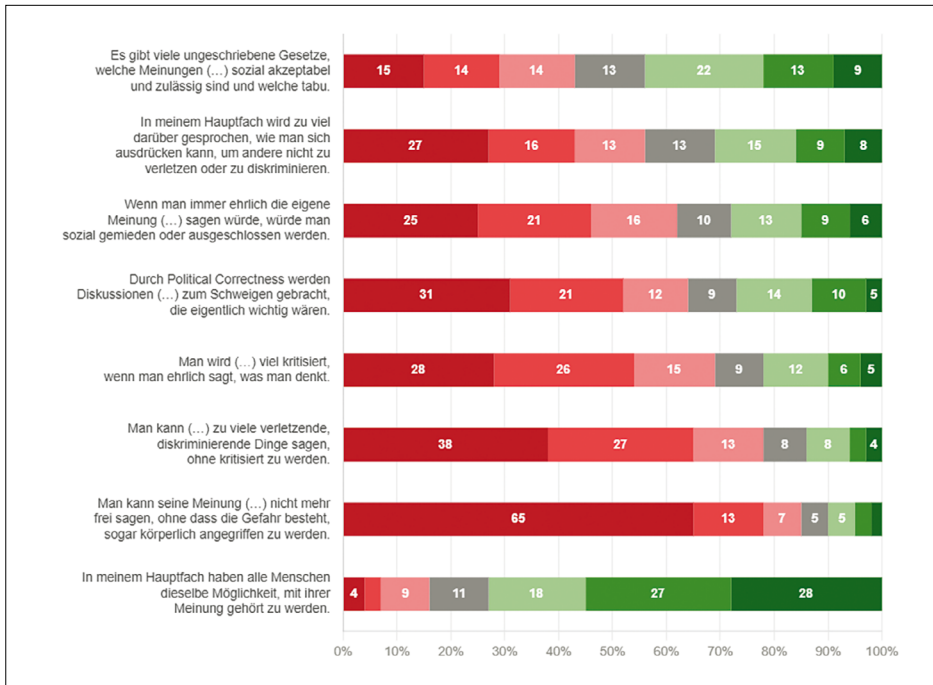
Die erste Forschungsfrage zielte auf die Frage ab, wie Studierende an deutschen Universitäten die Möglichkeiten wahrnehmen, sich frei zu äußern. Die Ergebnisse zeigen, dass eine große Mehrheit der Studierenden ein durchaus positives Bild von der Meinungsfreiheit an ihren Universitäten hat. Fast einhellig widersprechen die Studierenden (zumindest eher) der Aussage, dass die Gefahr eines körperlichen Angriffs bei einer Meinungsäußerung bestehe (85 %, $M = 1.91$, $SD = 1.58$), und fast ebenso einhellig wird verneint, dass „man zu viele verletzende, diskriminierende Dinge sagen [könne], ohne [dafür] kritisiert zu werden“ (78 %, $M = 2.46$, $SD = 1.67$). Fast ebenso viele sind nicht der Ansicht, dass man „zu viel kritisiert

[würde], wenn man ehrlich sagt, was man denkt“ (71 %, $M = 2.78$, $SD = 1.72$). Drei Viertel sind (zumindest eher) der Meinung, dass in ihrem Hauptfach alle Menschen dieselbe Möglichkeit haben, mit ihrer Meinung gehört zu werden (73 %, $M = 2.73$, $SD = 1.67$). Dass durch „Political Correctness“ Diskussionen unterdrückt würden, die eigentlich wichtig wären, wird von mehr als zwei Dritteln der Studierenden (zumindest eher) negiert (63 %, $M = 3.03$, $SD = 1.93$). Die Gefahr eines sozialen Ausschlusses, „wenn man immer ehrlich die eigene Meinung sagen würde“, sehen knapp zwei Drittel (eher) nicht (62 %, $M = 3.17$, $SD = 1.89$). Immerhin ein Viertel sieht diese Gefahr aber durchaus.

Etwas weniger einig sind sich die Studierenden bei zwei weiteren Items. So sind etwas mehr als die Hälfte nicht der Ansicht, dass in ihrem Hauptfach „zu viel darüber gesprochen [würde], wie man sich ausdrücken kann, um andere nicht zu verletzen oder zu diskriminieren“, während knapp ein Drittel durchaus diesen Eindruck hat ($M = 3.34$, $SD = 1.99$). Noch stärker polarisiert sind die Einschätzungen bei der Frage, ob es „viele ungeschriebene Gesetze darüber [gebe], welche Meinungen i[m] Hauptfach sozial akzeptabel und zulässig sind und welche tabu“ ($M = 3.86$, $SD = 1.89$). Hier stehen sich jeweils 43 % der Studierenden gegenüber, die dieser Aussage (eher) zustimmen, und die ihr (eher) nicht zustimmen. Möglicherweise hängt dies damit zusammen, dass sich die Aussage auf alle möglichen Arten von Meinungen beziehen lässt, sowohl solche mit einem inhaltlichen Fachbezug (etwa zu Theorien oder Methoden) als auch solche, die stärker politisch sind (Abbildung 1).

Unsere zweite Forschungsfrage zielte darauf ab, ob sich Studierende unterschiedlicher Fächer in ihrer Wahrnehmung der universitären Meinungsfreiheit unterscheiden. Eine Varianzanalyse zeigt keine signifikanten Differenzen ($F = 1.016$, $p = .407$). Betrachtet man die einzelnen Items, so ergeben sich zwar bei der Hälfte signifikante Differenzen zwischen einzelnen Fächergruppen, die allerdings überschaubar bleiben.

Abbildung 1. Wahrnehmung der Meinungsfreiheit in ihrem Hauptfach durch Studierende im Sommer 2023



Anmerkungen. $N = 845$ – 880 Befragte (ohne „weiß nicht“). 7-stufige Skala von „1“ – „stimme überhaupt nicht zu“ bis „7“ – „stimme voll und ganz zu“ (Ausweichkategorie „weiß nicht“). Frage: „Nun möchten wir Ihnen ein paar Fragen dazu stellen, wie Diskussionen in den Seminaren und Vorlesungen Ihres Hauptfachs ablaufen. Dazu finden Sie unten einige Aussagen. Stellen Sie sich vor, Sie sitzen in einem typischen Seminar oder einer typischen Vorlesung Ihres Hauptfachs. Wie sehr stimmen Sie folgenden Aussagen zu?“

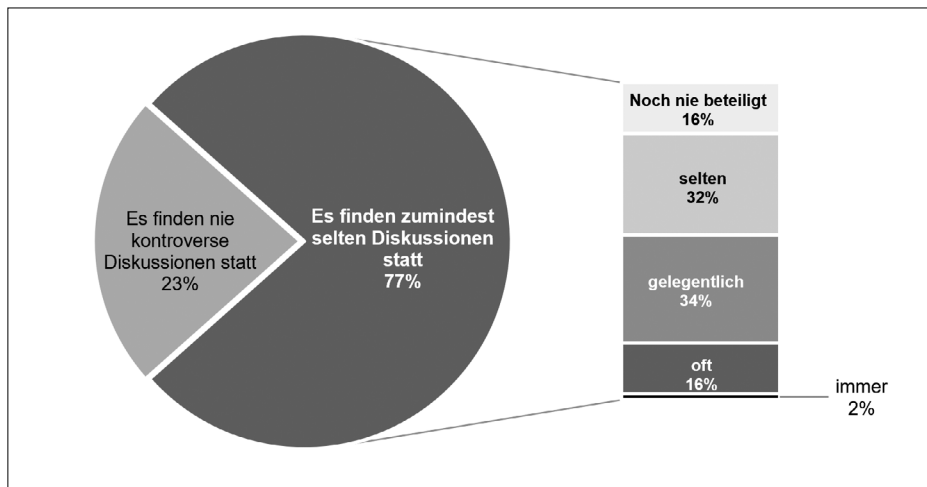
4.2 Diskussionen über kontroverse Themen und Redebereitschaft

Die dritte Forschungsfrage bezog sich darauf, wie häufig in Lehrveranstaltungen an deutschen Universitäten über aktuelle kontroverse Themen diskutiert wird und wie häufig sich Studierende an solchen Debatten beteiligen. Die Daten zeigen zunächst, dass Studierende im Durchschnitt nur gelegentlich solche Diskussionen wahrnehmen ($M = 2.27$, $SD = .94$, $n = 875$). Fast ein Viertel der Studierenden hat solche Diskussionen bislang sogar noch nie (23 %), fast zwei Fünftel (38 %) haben sie nur selten erlebt. Bei etwas mehr als einem Viertel der Studierenden war dies gelegentlich der Fall (27 %) und nur knapp 12 % geben an, dass in ihren Lehrveranstaltungen oft kontrovers diskutiert wird.

Diejenigen Studierenden, in deren Lehrveranstaltungen solche Diskussionen bereits stattfanden, haben sich zudem durchschnittlich nur selten bis gelegentlich daran beteiligt ($M = 2.55$, $SD = 1.00$, $n = 670$). Während fast die Hälfte der Stu-

dierenden noch nie oder selten mitdiskutiert hat (48 %), tun dies nur 18 % oft oder immer. 34 % gaben an, sich gelegentlich in solche Debatten einzubringen (Abbildung 2).

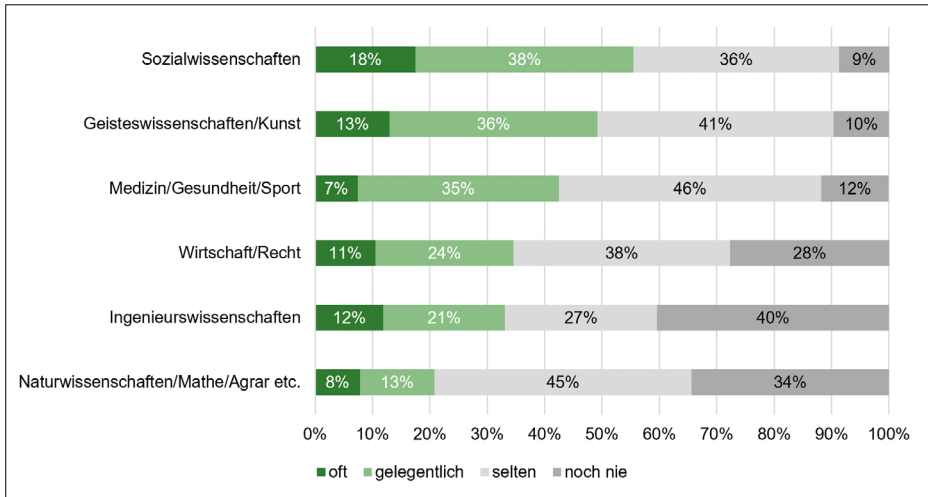
Abbildung 2. Häufigkeit von Diskussionen über kontroverse Themen und persönliche Beteiligung an solchen Diskussionen in den Lehrveranstaltungen des Hauptfachs im Sommer 2023



Anmerkungen. $N = 875$ (Häufigkeit von kontroversen Diskussionen) bzw. 670 Befragte (Häufigkeit der Redebereitschaft). Fragen: "Und wie häufig kam es in einer Vorlesung oder einem Seminar Ihres Hauptfachs bereits zu einer kontroversen Diskussion, zum Beispiel zu den bereits erwähnten Themen?"; "Wenn Sie an die eben angesprochenen Themen denken, wie häufig haben Sie sich selbst an Diskussionen in einem Seminar oder einer Vorlesung zu einem dieser Themen beteiligt?"

Die vierte Forschungsfrage fragte danach, ob sich Studierende unterschiedlicher Fächergruppen in Diskussionshäufigkeit und -beteiligung unterscheiden. Eine ein-faktorielle Varianzanalyse zeigt signifikante Unterschiede der Diskussionshäufigkeit je nach Fächergruppe (Welch $F(5, 384,356) = 14.322, p < .001, \eta^2 = .075, N = 875$). Die Fächergruppe kann hierbei 7,5 % der Varianz erklären. Der Post-hoc-Test nach Games-Howell zeigt, dass in den Lehrveranstaltungen der Mathematik, Natur- & Agrarwissenschaften ($M = 1.94, SD = .89$) signifikant seltener über kontroverse Themen diskutiert wird als in Medizin, Gesundheits- & Sportwissenschaften ($M = 2.37, SD = .79$), Geisteswissenschaften & Kunst ($M = 2.53, SD = .84$) sowie Sozialwissenschaften ($M = 2.64, SD = .88$). Auch Studierende der Ingenieurwissenschaften ($M = 2.04, SD = 1.05$) sowie der Rechts- & Wirtschaftswissenschaften ($M = 2.16, SD = .95$) nehmen weniger kontroverse Diskussionen wahr als Studierende der Geisteswissenschaften & Kunst sowie der Sozialwissenschaften (Abbildung 3).

Abbildung 3. Häufigkeit von kontroversen Diskussionen in den Lehrveranstaltungen verschiedener Fächergruppen im Sommer 2023



Anmerkungen. $N = 875$ (Häufigkeit von kontroversen Diskussionen). Frage: „Und wie häufig kam es in einer Vorlesung oder einem Seminar Ihres Hauptfachs bereits zu einer kontroversen Diskussion, zum Beispiel zu den bereits erwähnten Themen?“

Im Gegensatz dazu lassen sich bei der Redebereitschaft keine signifikanten Unterschiede zwischen den Fächergruppen feststellen ($M = 2.39$ bis $M = 2.57$; $F(5, 664) = 1.612$, $p = .154$, $N = 670$). Dies bedeutet: In den Fächergruppen wird zwar unterschiedlich häufig diskutiert, aber wenn, dann beteiligt sich ein ähnlicher Anteil Studierender.

4.3 Einflüsse auf die Wahrnehmung von Meinungsfreiheit

Um zu überprüfen, welche persönlichen Merkmale die Wahrnehmungen von Meinungsfreiheit beeinflussen, wurden lineare blockweise Regressionen durchgeführt. Das hier betrachtete Regressionsmodell 4 ist signifikant ($F(17, 818) = 22.017$, $p < .001$, $f^2 = .429$; $N = 836$) und erklärt 30,0 % der Varianz (korr. R^2). Die Robustness-Checks erbringen im Hinblick auf die Struktur der einflussreichen Prädiktoren keine wesentlich anderen Ergebnisse, einzelne Differenzen werden unten erwähnt (siehe auch Anhang D auf OSF).

Den größten Einfluss haben die allgemeinen politischen Prädispositionen der Studierenden. Sie tragen 13,6 % zur erklärten Varianz bei. Während die Politikverdrossenheit die Wahrnehmung der Meinungsfreiheit nicht signifikant beeinflusst, findet sich sowohl für die politische Orientierung ($\beta = .216$; $p < .001$) als auch (in allen Modellen mit Imputation) für die Demokratiezufriedenheit ($\beta = -.078$, $p = .021$) ein entsprechender Einfluss: Studierende, die sich auf der Links-Rechts-Skala weiter rechts einordnen, und solche, die unzufrieden mit der Demokratie sind, nehmen die Meinungsfreiheit auf dem Campus als einge-

schränkter wahr. Die politische Orientierung übt den stärksten Einfluss aller Merkmale aus.

Ebenfalls einen signifikanten Effekt haben Befürchtungen der Studierenden vor negativen Reaktionen ihres akademischen Umfelds. Dies zeigt sich sowohl für mögliche (soziale) Sanktionen durch Mitstudierende ($\beta = .112, p = .010$) als auch durch Lehrende ($\beta = .193, p < .001$). Je mehr die Studierenden Sanktionen erwarten, desto eingeschränkter empfinden sie die Meinungsfreiheit auf dem Campus. Der Block zu Sanktionsängsten erklärt zusätzlich 7,1 % der Varianz, und Befürchtungen im Hinblick auf die Lehrenden stellen den zweitstärksten Einflussfaktor dar.

Einen deutlich geringeren Einfluss haben Parteipräferenzen (Änderung korr. $R^2 = .032$). Hierbei zeigt sich nur für Bündnis 90/Die Grünen ein signifikanter Effekt ($\beta = -.136, p < .001$). Demnach schätzen Studierende, die die Grünen wählen würden, die Meinungsfreiheit auf dem Campus etwas positiver ein. Frühere Studien, die einen signifikant negativen Einfluss einer AfD-Präferenz auf die Wahrnehmung der Meinungsfreiheit in der Gesamtbevölkerung belegten, konnten in dieser Stichprobe nicht repliziert werden. Mögliche Ursachen dafür diskutieren wir unten.

Daneben spielen auch politische Kongruenz-Wahrnehmungen eine Rolle für die Einschätzung der Meinungsfreiheit. Sie tragen zusätzlich 2,5 % zur erklärten Varianz bei. Zumindest für die politische Kongruenz mit Mitstudierenden findet sich ein signifikanter Effekt, der nur in einzelnen Modellen mit Imputation und Gewichtung nicht erkennbar ist: Je größer die wahrgenommene Übereinstimmung, desto positiver wird die Meinungsfreiheit auf dem Campus beurteilt ($\beta = -.107, p = .004$). Die politische Kongruenz mit den Lehrenden spielt dagegen keine Rolle.

Schließlich ergeben sich auch signifikante Effekte der Kontrollvariablen: Studierende mit höherem politischen Interesse ($\beta = .159, p < .001$) und männliche Studierende nehmen die Meinungsfreiheit auf dem Campus negativer wahr ($\beta = -.082, p < .010$). Das Alter hat dagegen keinen signifikanten Effekt. Die Kontrollvariablen erklären insgesamt 5 % der Varianz.

Die Tabelle basiert auf Modell 4 mit Gewichtung und Mittelwert-Imputation. Der Effekt der Demokratiezufriedenheit ist in allen Modellen mit Imputation signifikant, der Effekt der politischen Kongruenz mit Mitstudierenden in fünf Modellen mit Gewichtung und Imputation nicht signifikant (Anhang D auf OSF).

Tabelle 1. Einflüsse auf die Wahrnehmung universitärer Meinungsfreiheit durch Studierende (Modell 4)

Regressionskoeffizienten	β	SE	KI (95 %)		p
			UG	OG	
<i>Kontrollvariablen</i>					
Alter	-.019	.007	-.019	.010	.520
Geschlecht**	-.082	.076	-.347	-.048	.010
Politisches Interesse***	0.159	.030	.061	.178	<.001
Änderung in R ²	.050				
<i>Block 1: Politische Prädispositionen</i>					
Politische Orientierung***	.216	.026	.091	.193	<.001
Politikverdrossenheit	.010	.052	-.090	.114	.818
Demokratiezufriedenheit*	-.078	.062	-.265	-.021	.021
Änderung in R ²	.136				
<i>Block 2: Parteipräferenz</i>					
Parteipräferenz: CDU/CSU	-.010	.139	-.311	.233	.778
Parteipräferenz: Grüne***	-.136	.105	-.564	-.151	<.001
Parteipräferenz: AfD	.032	.215	-.230	.614	.371
Parteipräferenz: SPD	.009	.126	-.215	.278	.799
Parteipräferenz: FDP	.003	.158	-.296	.323	.934
Parteipräferenz: Die Linke	-.013	.158	-.369	.251	.709
Parteipräferenz: Nicht-Wahl	-.050	.204	-.724	.078	.114
Änderung in R ²	.032				
<i>Block 3: Politische Kongruenz</i>					
Kongruenz Studierende**	-.107	.032	-.154	-.029	.004
Kongruenz Lehrende	-.047	.035	-.115	.025	.205
Änderung in R ²	.025				
<i>Block 4: Sanktionen infolge von Meinungsäußerungen</i>					
Sanktionsangst Mitstudierende**	.112	.037	.023	.169	.010
Sanktionsangst Lehrende***	.193	.034	.087	.219	<.001
Änderung in R ²	.071				
R ² (korr.)	.300	F(17, 818) = 22.017			<.001
Effektstärke f ²	.429				

Anmerkungen. KI = Konfidenzintervall; UG = Untergrenze; OG = Obergrenze. Geschlecht: 1 = weiblich. Politische Orientierung: 1 = links bis 11 = rechts. $N = 843$. *** $p < .001$; ** $p < .01$; * $p \leq .05$.

4.4 Einflüsse auf die Beteiligung an Diskussionen über kontroverse Themen

Um zu prüfen, welche individuellen Merkmale die Redebereitschaft der Studierenden beeinflussen, wurden die gleichen Prädiktoren wie bei den vorherigen Analysen in die Modelle aufgenommen. Das hier betrachtete Regressionsmodell D ist als Ganzes signifikant ($F(17,613) = 8.494, p < .001, f^2 = .195$) und erklärt mit insgesamt 16,8 % (korr. R^2) deutlich weniger Varianz als die Modelle zur Wahrnehmung der Meinungsfreiheit. Aufgrund fehlender Werte basieren die Ergebnisse auf einer Fallzahl von $N = 613$ (Tabelle 2). Die Robustness-Checks erbringen im Hinblick auf die Struktur der einflussreichen Prädiktoren einzelne Differenzen, die unten erwähnt werden (siehe auch Anhang D auf OSF).

Politische Prädispositionen erklären nur 3 % der Varianz (Änderung in R^2). Dabei hat die Politikverdrossenheit einen signifikant negativen Einfluss und erweist sich als stärkster Einflussfaktor: Politikverdrossene Studierende beteiligen sich demnach seltener an Diskussionen über kontroverse Themen ($\beta = -.201, p < .001$). Rechts-Links-Orientierung und Demokratiezufriedenheit haben dagegen anders als bei der Wahrnehmung der Meinungsfreiheit keinen signifikanten Effekt.

Die Parteipräferenzen leisten mit 3,1 % ebenfalls einen geringen Beitrag zur Varianzaufklärung. Dabei zeigt sich, dass Studierende mit einer Präferenz für die AfD eine signifikant höhere Redebereitschaft aufweisen ($\beta = .137, p = .003$). Dies gilt auch für Studierende, die die SPD ($\beta = .123, p = .008$) oder die Linke ($\beta = .090, p = .044$) unterstützen. Allerdings sind die letztgenannten Effekte nicht völlig robust. Bei den übrigen Parteien zeigt sich kein Einfluss auf die Redebereitschaft.

Anders als bei der Wahrnehmung der Meinungsfreiheit haben die politische Kongruenz mit Mitstudierenden und Lehrenden keinen signifikanten Einfluss auf die Redebereitschaft (Änderung $R^2 = .001$; Mitstudierenden: $\beta = -.035, p = .469$; Lehrenden: $\beta = -.007, p = .892$). Als einflussreich und völlig robust erweisen sich allerdings die Sanktionsängste gegenüber Mitstudierenden. Studierende, die mehr Sanktionen durch ihre Kommilitonen befürchten, beteiligen sich seltener an universitären Diskussionen über kontroverse Themen ($\beta = -.152, p = .005$). Dagegen sind die Befunde für die befürchteten Sanktionen durch Lehrende nicht über alle Modelle hinweg stabil. Wie im hier fokussierten Modell ($\beta = .086, p = .102$) zeigt sich in den Modellen mit Gewichtung durchweg kein signifikanter Einfluss und in den Modellen ohne Gewichtung auch nicht in allen Fällen. Es handelt sich also nicht um einen robusten Befund.

Als einflussreich erweisen sich zudem die Kontrollvariablen, die weitere 11,5 % der Varianz erklären: Je höher das politische Interesse, desto eher beteiligen sich Studierende an Diskussionen ($\beta = .185, p < .001$). Für das Geschlecht konnten keine Effekte nachgewiesen werden, das Alter übt nur in Modellen ohne Imputation einen Einfluss aus, der deshalb hier nicht als robust angesehen und nicht weiter interpretiert wird.

Die meisten der beschriebenen Befunde sind robust: Die signifikanten Zusammenhänge bleiben sowohl in den ungewichteten Daten als auch unabhängig von den Mittelwertimputationen bestehen. Insbesondere die Politikverdrossenheit, die Unterstützung der AfD, die Kongruenz mit Mitstudierenden und das politische Interesse konnten in den verschiedenen Modellen als signifikante Prädiktoren identifiziert werden. Der Einfluss der Präferenz für die SPD ist über die Modelle hinweg weniger stabil und sollte deshalb vorsichtiger interpretiert werden.

Tabelle 2. Einflüsse auf die persönliche Redebereitschaft Studierender (Modell D)

Regressionskoeffizienten	β	SE	KI (95 %)		p
			UG	OG	
Kontrollvariablen					
Alter	.072	.008	.000	.031	.057
Geschlecht	.071	.081	-.016	.301	.079
Politisches Interesse***	.185	.032	.056	.181	<.001
Änderung in R^2	.119				
<i>Block 1: Politische Prädispositionen</i>					
Politische Orientierung	.052	.027	-.026	.080	.315
Politikverdrossenheit***	-.201	.053	-.309	-.102	<.001
Demokratiezufriedenheit	.002	.062	-.120	.125	.964
Änderung in R^2	.030				
<i>Block 2: Parteipräferenz</i>					
Parteipräferenz: CDU/CSU	-.004	.142	-.292	.265	.924
Parteipräferenz: Grüne	.010	.112	-.197	.243	.838
Parteipräferenz: AfD**	.137	.211	.210	1.040	.003
Parteipräferenz: SPD**	.123	.130	.090	.600	.008
Parteipräferenz: FDP	.049	.163	-.136	.503	.260
Parteipräferenz: Die Linke*	.090	.157	.009	.627	.044
Parteipräferenz: Nicht-Wahl	-.026	.229	-.602	.296	.504
Änderung in R^2	.031				
<i>Block 3: Politische Kongruenz</i>					
Kongruenz Studierende	-.035	.034	-.090	.042	.469
Kongruenz Lehrende	-.007	.036	-.076	.066	.892
Änderung in R^2	.001				
<i>Block 4: Sanktionen infolge von Meinungsäußerungen</i>					
Sanktionsangst Mitstudierende**	-.152	.037	-.176	-.032	.005
Sanktionsangst Lehrende	.086	.032	-.010	.117	.102
Änderung in R^2	.011				
R^2 (korr.)	.168	$F(17, 613) = 8.494$			<.001
Effektstärke f^2	.202				

Anmerkungen. KI = Konfidenzintervall; UG = Untergrenze; OG = Obergrenze. Geschlecht: 1 = weiblich. Politische Orientierung: 1 = links bis 11 = rechts. $N = 631$. *** $p < .001$; ** $p < .01$; * $p \leq .05$.

Die Tabelle basiert auf Modell D mit Gewichtung und Mittelwert-Imputation. Der Effekt der Parteipräferenz für die SPD ist nur in Modellen mit Imputation und Gewichtung, der Effekt der Parteipräferenz für die Linke nur in Modellen mit Imputation und der Effekt des Alters nur in Modellen ohne Imputation signifikant. Effekte der Sanktionsangst vor Lehrenden waren nur in Modellen ohne Gewichtung festzustellen (Anhang D auf OSF).

5. Diskussion und Limitationen

Wie nehmen Studierende die Situation der Meinungsfreiheit an deutschen Universitäten wahr? Wie intensiv beteiligen sie sich dort an Diskussionen über kontroverse Themen? Und welchen Einfluss haben Merkmale wie politische Einstellungen, politische Kongruenz und die Furcht vor den Reaktionen von Mitstudierenden und Lehrenden auf ihre Wahrnehmungen und ihr Verhalten? Vor dem Hintergrund einer intensiven öffentlichen Debatte sowie erster, kontrovers diskutierter Befunde (Revers & Traunmüller, 2020; Traunmüller & Revers, 2021) beschäftigte sich die vorliegende Studie mit diesen bislang nicht auf breiter Basis untersuchten Fragen. Die Befunde lassen sich wie folgt zusammenfassen:

(1) Die weit überwiegende Mehrheit der Studierenden, je nach konkreter Fragestellung etwa zwei Drittel bis drei Viertel, sieht die Meinungsfreiheit auf dem Campus durchaus gewährleistet. Gleichzeitig nimmt die überwiegende Mehrheit ein Klima wahr, in dem Diskriminierung nicht toleriert wird. Demgegenüber steht eine Minderheit von – je nach konkreter Fragestellung – etwa einem Fünftel der Studierenden, die die Meinungsfreiheit in ihrem Hauptfach kritischer sehen. Zudem erkennen viele Studierende „ungeschriebene Gesetze“, gegen die man nicht verstoßen sollte. Allerdings lässt sich aufgrund der Fragestellungen nicht genauer konkretisieren, ob sich die „ungeschriebenen Gesetze“ auf wissenschaftliche oder gesellschaftliche bzw. politische Sachverhalte beziehen.

(2) Im Hinblick auf die Wahrnehmung universitärer Meinungsfreiheit zeigen sich keine Differenzen zwischen den Studierenden verschiedener Fächergruppen. Dies widerspricht der Annahme (z. B. Revers & Traunmüller, 2020), dass die Meinungsfreiheit in Fächern als eingeschränkter wahrgenommen wird, in denen häufiger über kontroverse gesellschaftliche Themen diskutiert wird, wie etwa in den Sozialwissenschaften (siehe unten).

(3) Trotz des Anspruchs von Universitäten, Orte des freien Austauschs über gesellschaftlich relevante Themen zu sein, kommt es nach Einschätzung der Studierenden im Durchschnitt aller Fächergruppen nur ab und zu zu kontroversen Diskussionen in Lehrveranstaltungen. Allerdings zeigen sich deutliche Unterschiede zwischen den Fächergruppen: So finden in Mathematik, den Natur- & Agrarwissenschaften, den Ingenieurwissenschaften sowie den Rechts- & Wirtschaftswissenschaften seltener kontroverse Debatten statt als in Geisteswissenschaften & Kunst sowie den Sozialwissenschaften. Hinsichtlich der Redebereitschaft unterscheiden sich die Fächergruppen jedoch kaum voneinander: Die meisten Studierenden sind eher zurückhaltend. Während knapp jeder zweite Studierende bei kontroversen Diskussionen – so sie denn stattfinden – immer oder fast immer schweigt, bringt sich nur knapp ein Fünftel oft oder immer ein.

(4) Die Einflüsse auf die beiden hier im Zentrum stehenden Aspekte akademischer Meinungsfreiheit fallen – anders als erwartet – recht unterschiedlich aus. Zum einen lässt sich die Wahrnehmung der universitären Meinungsfreiheit durch die einbezogenen Merkmale der Studierenden sehr viel besser erklären als die individuelle Redebereitschaft. Das bedeutet, dass die Beteiligung an Diskussionen in Seminaren und Vorlesungen in noch stärkerem Maße durch Charakteristika der Studierenden geprägt wird, die wir hier nicht berücksichtigt haben. Zum anderen

unterscheiden sich auch die Muster der Einflüsse, was darauf hindeutet, dass man die Wahrnehmung universitärer Meinungsfreiheit nicht mit der individuellen Redebereitschaft gleichsetzen darf.

(5) Was die Erklärungsfaktoren angeht, haben sich unsere Annahmen nur zum Teil bestätigt. So wird die Wahrnehmung der Meinungsfreiheit in erster Linie durch politische Einstellungen beeinflusst, darunter die politische Orientierung, die Demokratiezufriedenheit und die Parteipräferenz. Während sich die Mehrheit der Studierenden eher im politisch linken Spektrum bewegt, nehmen Studierende, die sich politisch weiter rechts verorten, stärkere Einschränkungen wahr. Dies entspricht den Befunden von Rothut et al. (2023) und Unzicker (2018) für die Gesamtbevölkerung in Deutschland. Stärkere Einschränkungen werden auch von Studierenden wahrgenommen, die unzufriedener mit der Demokratie sind, womit sich der von Quiring et al. (2020) festgestellte Zusammenhang zwischen wahrgenommenen Sprechverboten und Demokratiezufriedenheit auf den Campus übertragen lässt. Zum anderen nehmen Studierende mit einer Präferenz für Bündnis 90/Die Grünen weniger Einschränkungen wahr, wohingegen sich Präferenzen für andere Parteien nicht systematisch auswirken – auch nicht im Fall der AfD. Daneben führt, wie angenommen, die Furcht vor negativen Reaktionen von Lehrenden und Mitstudierenden auf eigene Meinungsäußerungen zu einer verstärkten Wahrnehmung eingeschränkter Meinungsfreiheit. Und schließlich wirkt sich auch eine stärkere politische Kongruenz bzw. Inkongruenz mit dem akademischen Umfeld auf die wahrgenommene Meinungsfreiheit aus – allerdings nur im Hinblick auf Mitstudierende: Nehmen Studierende hier eine höhere politische Ähnlichkeit wahr, sehen sie auch weniger Probleme mit der Meinungsfreiheit. Für die Lehrenden zeigt sich dieser Einfluss hingegen nicht.

Bemerkenswerterweise spielt Politikverdrossenheit, die Rothut et al. (2023) als stärksten Prädiktor für die wahrgenommene Meinungsfreiheit identifizierten, im Falle der Studierenden und ihrer Beurteilung der Meinungsfreiheit auf dem Campus keine signifikante Rolle. Hinsichtlich der Kontrollvariablen zeigt sich, dass politisch interessierte Studierende die Meinungsfreiheit als eingeschränkter wahrnehmen (ähnlich bei Rothut et al., 2020). Möglicherweise führen sie häufiger Gespräche über politische Themen und werden öfter mit konträren Positionen konfrontiert.

(6) Die aktive Teilnahme an kontroversen Diskussionen lässt sich nur zum Teil durch die gleichen Faktoren erklären wie die Wahrnehmung der universitären Meinungsfreiheit. Der stärkste Einfluss geht hier von der Politikverdrossenheit aus, wobei sich politikverdrossenere Studierende seltener an Diskussionen beteiligen. Vermutlich sind sie eher resigniert und sehen weniger Sinn darin, sich an politischen Debatten zu beteiligen. Demgegenüber beteiligen sich politisch interessierte Studierende häufiger an kontroversen Diskussionen. Dies könnte auf eine höhere Motivation oder fundiertere Positionen zurückzuführen sein, die hier nicht weiterführend untersucht werden konnten. Anders als bei der Wahrnehmung der Meinungsfreiheit haben die politische Orientierung, die Demokratiezufriedenheit, die politische Kongruenz sowie das Geschlecht keinen Einfluss auf die Redebereitschaft. Allerdings reduzieren Ängste vor Sanktionen durch Mitstudierende die Bereitschaft, an Diskussionen teilzunehmen. Ängste vor Sanktionen durch Lehrende

spielen in diesem Kontext dagegen keine Rolle. Damit kann eine der zentralen Annahmen der Theorie der Schweigespirale (Noelle-Neumann, 1974), zumindest in Teilen, gestützt und bisherige vergleichbare Befunde (Matthes et al., 2018; Neubaum & Krämer, 2018) auf den Universitätskontext ausgeweitet werden.

Etwas überraschender fällt der Einfluss der Parteipräferenzen aus: Studierende mit Präferenzen für AfD – und tendenziell auch für SPD und Linke –, die also gerade nicht zur stärksten Gruppe der Bündnis 90/Die Grünen-Anhängerschaft gehören, beteiligen sich nach eigenen Angaben besonders häufig an kontroversen Debatten. Dies könnte dafür sprechen, dass diese Studierenden ein besonderes Interesse haben, dass auch ihre Positionen in den Diskussionen vertreten sind, und es sich bei ihren Meinungsäußerungen in diesem Sinne um „corrective action“ (Wintterlin et al., 2021) handelt, die von Mitgliedern eines „harten Kerns“ und „Avantgardisten“ (Noelle-Neumann, 1980, S. 200) bzw. eines „fish-out-of-water“ (Norris, 2023) ausgeübt wird. Somit liegt der Schluss nahe, dass Meinungsäußerung dann nicht trotz, sondern gerade wegen der Wahrnehmung stattfindet, dass man sich in der Minderheit befinde. Allerdings gilt es hier zu berücksichtigen, dass die politische Kongruenz als solche keinen signifikanten Einfluss auf die Redebereitschaft hat. Es ist also nicht eindeutig, ob die höhere Beteiligung von Studierenden, die mit AfD, SPD oder Der Linken sympathisieren, tatsächlich durch wahrgenommene Mehr- oder Minderheitspositionen zu erklären ist. Zudem handelt es sich um schwache Effekte, weshalb der Einfluss der Parteipräferenzen nicht überinterpretiert werden sollte.

Die recht geringe Erklärungskraft der politischen Prädiktoren lässt mögliche Schweigespiraleffekte unwahrscheinlich erscheinen und den Schluss zu, dass für die Redebereitschaft stabilere individuelle Prädispositionen eine entscheidendere Rolle spielen. Darunter könnten Persönlichkeitsmerkmale wie Extrovertiertheit und Selbstbewusstsein fallen. Beteiligen sich Studierende selten an Diskussionen, könnte dies unter anderem auch daran liegen, dass sie insgesamt dazu neigen, ihre eigene Meinung tendenziell eher zurückzuhalten (Hayes et al., 2005) oder Angst davor haben, falsch zu liegen (Stevens & Haidt, 2018b). Eine stärkere Beteiligung könnte hingegen mit einer hohen Motivation oder Leistungsorientierung assoziiert sein, wenn man annimmt, dass Studierende glauben, auch mit Beiträgen zu kontroversen Themen einen guten Eindruck bei Lehrenden und/oder Mitstudierenden hinterlassen zu können.

Die Befunde ermöglichen nach unserem Kenntnisstand erstmals Einblicke in die wahrgenommene Meinungsfreiheit sowie Redebereitschaft an deutschen Universitäten aus Studierendenperspektive und zeigen relevante Einflussfaktoren auf. Dennoch unterliegt die vorliegende Studie einigen Limitationen. So fand die Datenerhebung über ein Online-Access-Panel statt; es handelt sich also nicht um eine Zufallsstichprobe. Die Selbstselektion der Teilnehmenden hat vermutlich zu Über- bzw. Unterrepräsentation bestimmter Gruppen von Studierenden beigetragen (Taddicken, 2013). Auch ist denkbar, dass Studierende, die eine hohe Politikverdrossenheit aufweisen oder sich an den politischen Extremen verorten, nicht Teil des Panels sind (Merkley, 2020). Da eine echte Zufallsstichprobe nicht mit vertretbarem Aufwand herzustellen war, eine aufwändige Gewichtung realisiert wurde und die Stichprobe in vielen weiteren Merkmalen (z. B. regionale Verteilung,

politische Einstellungen auch nach Fächern) die Grundgesamtheit gut zu repräsentieren scheint, sind wir dennoch davon überzeugt, dass unsere Analyse vor allem im Hinblick auf die Zusammenhangsanalysen aussagekräftige Ergebnisse liefert.

Eine weitere Limitation besteht darin, dass bei einigen für die Analyse wichtigen Merkmalen eine nicht unerhebliche Zahl fehlender Werte zu beobachten war. Obwohl die Ursachen dafür nicht ohne Weiteres zu klären sind, kann man vermuten, dass sich manche Studierende hier schlicht kein Urteil zutrauten – etwa über die politischen Einstellungen von Mitstudierenden und Lehrenden. Ebenfalls kann man vermuten, dass dies vor allem bei Studierenden in bestimmten Fächern oder mit bestimmten weiteren Merkmalen der Fall war (z. B. Studierenden mit geringem politischem Interesse oder Wissen). Allerdings stand die Aufklärung dieser Fragen nicht im Mittelpunkt des Beitrags und wir haben möglichen Effekten auf die Analysen mit unseren Robustness-Checks umfassend Rechnung getragen.

Des Weiteren sind trotz einer recht hohen Fallzahl manche analytisch interessanten Gruppen sehr klein. Dies trifft etwa für die Studierenden mit einer Präferenz für die AfD zu ($n = 29$). Zudem handelt es sich um eine Querschnittsbefragung, sodass Aussagen über die kausale Richtung von Zusammenhängen auf (plausiblen) theoretischen Annahmen basieren und nicht qua Untersuchungsdesigns getroffen werden können. Außerdem bleibt zu betonen, dass die vorgenommene Gewichtung Verzerrungen zwar reduzieren kann, aber Verzerrungen in den Daten vorliegen können (Arzheimer, 2009; Sand & Kunz, 2020).

Hinsichtlich der abhängigen Konstrukte muss darauf hingewiesen werden, dass sich die Frage zur Redebereitschaft lediglich auf Lehrveranstaltungen bezieht und damit Diskussionen, die darüber hinaus auf dem Campus stattfinden, nicht miteinschließt. Zudem wurde die Toleranz Studierender gegenüber spezifischen Situationen und Positionen im Hochschulkontext, die ebenfalls erhoben wurde, aufgrund der inhaltlichen Fokussierung des Beitrags nicht berücksichtigt (dazu Diehl et al., 2025). Des Weiteren geht diese Studie weder auf den Zusammenhang zwischen der Wahrnehmung der Meinungsfreiheit und der Redebereitschaft noch auf die Kausalrichtungen zwischen diesen Konstrukten und den einzelnen Prädiktoren ein. Zum Beispiel bleibt offen, ob das Gefühl politischer Inkongruenz zu Mitstudierenden dazu führt, die Meinungsfreiheit als eingeschränkter zu betrachten, oder erst die Wahrnehmung von Einschränkungen auf dem Campus das Gefühl von politischer Inkongruenz hervorruft. Ebenso wie Rothut et al. (2023) verweisen auch wir auf die Notwendigkeit längsschnittlicher (Panel-)Studien in diesem Bereich.

Hinsichtlich der Angst vor Sanktionen ist anzumerken, dass der hier verwendete Summenindex ein eher grober Indikator für mögliche Effekte befürchteter Sanktionen ist. Denn weder können wir differenzieren, wie häufig die jeweiligen Sanktionen befürchtet werden, noch wie groß die Befürchtungen sind, die diese auslösen. Wir können deshalb auch nicht differenzieren, welche möglichen Sanktionen als besonders schlimm empfunden werden und sich deshalb möglicherweise besonders stark auf die Wahrnehmung der Meinungsfreiheit und die Redebereitschaft auswirken.

Zuletzt bleibt anzumerken, dass neben individuellen Merkmalen auch weitere Faktoren die Wahrnehmung der Meinungsfreiheit auf dem Campus beeinflussen können (Norris, 2023; Rothut et al., 2023). Beispiele hierfür sind medial vermittelte Eindrücke (z. B. Meltzer, 2017) sowie institutionelle oder gesellschaftliche Strukturen, die das Urteil über die Meinungsfreiheit an Universitäten ebenfalls prägen könnten. Weiterführende Forschung sollte das Thema daher nicht nur auf Individualebene betrachten, sondern auch Meso- und Makrofaktoren sowie alternative theoretische Perspektiven einbeziehen, um ein umfassenderes Verständnis der Meinungsfreiheit auf dem Campus zu erlangen.

Diese Studie leistet trotz der genannten Einschränkungen einen Beitrag zur Debatte um Meinungsfreiheit an deutschen Universitäten. Gerade im Kontext des Nahostkonflikts bleibt dieses Thema relevant (dazu etwa Amitay et al., 2025; Eikmanns & Çay 2024; Grimm et al., 2025; Schieritz & Scholz, 2024). Vor dem Hintergrund aktueller Debatten (Eppelsheim, 2022; Mansour, 2023; ZEIT-Stiftung, 2024) und der Befunde dieser Studie wird deutlich, dass einerseits die empirische Untersuchung wahrgenommener Meinungsfreiheit an Universitäten von zentraler Bedeutung ist und andererseits das Führen konstruktiver universitärer Debatten aktiv gefördert werden sollte (Haake, 2021). Eine wahrgenommene Einschränkung der Meinungsfreiheit sowie geringe Redebereitschaft aufgrund befürchteter Sanktionen könnten zur Folge haben, dass intellektueller Fortschritt verhindert wird (Zhou & Barbaro, 2023) und sich Studierende an Universitäten nicht mehr wohl und zugehörig fühlen. Erste Ansatzpunkte zur Verbesserung der Debattenkultur liefern in diesem Zusammenhang Scott-Baumann und Perfect (2021). Auch qualitative Forschungsdesigns könnten weiteren Aufschluss darüber geben, warum Menschen sich (nicht) an politischen Diskussionen beteiligen und was sie überhaupt unter Meinungsfreiheit verstehen. Daraus können Maßnahmen zur Förderung einer akademischen Debattenkultur abgeleitet werden, die Universitäten als zentrale Orte gesellschaftlichen Diskurses neu positionieren.

Erklärung zum Einsatz generativer KI

(Generative) Künstliche Intelligenz wurde bei der Erstellung des Manuskripts oder der statistischen Analyse nicht verwendet. Allein Abbildung 1 wurde mit claude.ai grafisch überarbeitet.

Anhang

Der Anhang auf OSF ist über den folgenden Link abrufbar: <https://osf.io/cwkggt/overview>.

Literaturverzeichnis

- Amitay, G., Asraf, K., Elisha, E., Farber, S., Peleg-Koriat, I., & Yaron-Antar, A. (2025). Freedom of expression in Israeli campuses and social media during the war with Hamas. *Journal of Higher Education Policy and Management*, 48(1), 37–56. <https://doi.org/10.1080/1360080X.2025.2535028>

- Armingeon, K. (2001). *Fachkulturen, soziale Lage und politische Einstellungen der Studierenden der Universität Bern*. Universität Bern.
- Arzheimer, K. (2009). Gewichtungsvariation. In H. Schoen, H. Rattinger, & O. Gabriel (Hrsg.), *Vom Interview zur Analyse* (S. 361–388). Nomos.
- Barendt, E. (2010). *Academic freedom and the law: A comparative study*. Hart Publishing.
- Bar-Tal, D. (2017). Self-censorship as a socio-political-psychological phenomenon: Conception and research. *Advances in Political Psychology*, 38(S1), 37–65. <https://doi.org/10.1111/pops.12391>
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Beierlein, C., Kemper, C. J., Kovaleva, A., & Rammstedt, B. (2014). *Political efficacy Kurzsкала (PEKS)*. ZIS. Open Access Repositorium für Messinstrumente. GESIS Leibniz Institut für Sozialwissenschaften. <https://doi.org/10.6102/ZIS34>
- Berman, R. A. (2022). Inklusivität und die Grenzen der Wissenschaftsfreiheit: Zur Wiederkehr der »repressiven Toleranz«. In S. Kostner (Hg.), *Wissenschaftsfreiheit. Warum dieses Grundrecht zunehmend umkämpft ist* (S. 221–241). Nomos. <https://doi.org/10.5771/9783748928058>
- Breuer, I. (2019, 21. Februar). Meinungsfreiheit. Wo die Grenzen des Sagbaren liegen [Audio-podcast-episode]. In *Deutschlandfunk – Aus Kultur- und Sozialwissenschaften*. Deutschlandfunk. <https://www.deutschlandfunk.de/meinungsfreiheit-wo-die-grenzen-des-sagbaren-liegen-100.html>
- Burnett, A., Knighton, D., & Wilson, C. (2022). The self-censoring majority: How political Identity and ideology impacts willingness to self-censor and fear of isolation in the United States. *Social Media + Society*, 8(3). <https://doi.org/10.1177/20563051221123031>
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591–621. <https://doi.org/10.1146/annurev.psych.55.090902.142015>
- Cook, P., & Heilmann, C. (2013). Two types of self-censorship: Public and private. *Political Studies*, 61(1), 178–196. <https://doi.org/10.1111/j.1467-9248.2012.00957.x>
- Decker, F., Best, V., Fischer, S., & Küppers, A. (2019). *Vertrauen in Demokratie. Wie zufrieden sind die Menschen in Deutschland mit Regierung, Staat und Politik?* Friedrich-Ebert-Stiftung. <https://library.fes.de/pdf-files/fes/15621-20190822.pdf>
- Diehl, C., Revers, M., Traunmüller, R., Weidmann, N., & Wuttke, A. (2025). Students' motives for restricting academic freedom: Viewpoint discrimination and prosocial concerns. *PNAS*, 122(47). <https://doi.org/10.1073/pnas.2503804122>
- Döbele, C., Engels, J. N., Heinrich, R., Loew, N., Schläger, C., Simon, A. M., & Vitt, A.-K. (2023). *Krisenerwachen. Wie blicken junge Wähler:innen auf Politik, Parteien und Gesellschaft?* Friedrich-Ebert-Stiftung. <https://library.fes.de/pdf-files/a-p-b/20355.pdf>
- Eikmanns, F., & Çay, Y. E. (2024, 30. Mai). Propalästina-Proteste an deutschen Unis. Wo endet die Meinungsfreiheit? *taz*. Abgerufen am 21. August 2024 von <https://taz.de/Propalaestina-Proteste-an-deutschen-Unis!/6012172/>
- Eppelsheim, P. (2022, 3. Juli). Cancel Culture an der Uni. Der Gesinnungsterror linker Aktivisten. *Frankfurter Allgemeine Zeitung*. Abgerufen am 3. Januar 2024 von <https://www.faz.net/aktuell/politik/inland/linke-aktivisten-verhindern-vortrag-ueber-gen-der-an-humboldt-uni-18147133.html>
- Erich, N. (2020, 12. November). Meinungsfreiheit an Universitäten. Die Uni ist nichts für Schneeflockchen. *Zeit Online*. Abgerufen am 3. Januar 2024 von <https://www.zeit.de/kultur/2020-11/meinungsfreiheit-universitaeten-studie-cancel-culture-gesellschaft>

- Eßer, A. (2023). *Studentische Fachkulturen: Lebensstile und politische Dispositionen. Eine Untersuchung der Studienfächer Rechtswissenschaft, VWL, BWL, Sozialwissenschaften, Philosophie, Mathematik und Biologie*. Barbara Budrich.
- Freedom House (2023). *Freedom in the world 2023. Germany*. Freedom House. Abgerufen am 11. Januar 2024 von <https://freedomhouse.org/country/germany/freedom-world/2023>
- GESIS (o. D.). *Links-Rechts Orientierung*. GESIS Leibniz-Institut für Sozialwissenschaften. Abgerufen am 28. Januar 2024 von <https://www.gesis.org/angebot/daten-aufbereiten-und-analysieren/question-link/links-rechts-orientierung>
- Garton Ash, T. (2016). *Free speech: Ten principles for a connected world*. Yale University Press.
- Gibson, J. L., & Sutherland, J. L. (2023). Keeping your mouth shut: Spiraling self-censorship in the United States. *Political Science Quarterly*, 138(3), 361–376. <https://doi.org/10.1093/psquar/qqad037>
- Grewenig, E., Lergetporer, P., Simon, L., Werner, K., & Woessmann, L. (2023). Can internet surveys represent the entire population? A practitioners' analysis. *European Journal of Political Economy*, 78. <https://doi.org/10.1016/j.ejpoleco.2023.102382>
- Grimm, J., Chojnacki, S., Moya Schreieder, N., El Ghoubashy, I., & Sixta, T. (2025). Deutsche Wissenschaft seit dem 7. Oktober: Selbstzensur und Einschränkungen unter Forschenden mit Nahostbezug. *Working Papers Peace & Conflict Research*, 2. Freie Universität Berlin. https://www.interact.fu-berlin.de/_media/WP202501_v0_7.pdf
- Haake, K. (2021). Meinungs- und Lehrfreiheit – Was müssen Hochschulen aushalten? Bericht über die Tagung des Vereins zur Förderung des deutschen und internationalen Wissenschaftsrechts e.V. am 24.06.2021. *Ordnung der Wissenschaft*, 4, 257–264.
- Hayes, A. F., Glynn, C. J., & Shanahan, J. (2005). Validating the willingness to self-censor scale: Individual differences in the effect of the climate of opinion expression. *International Journal of Public Opinion Research*, 17(4), 443–455. <https://doi.org/10.1093/ijpor/edh072>
- Heil, L. U. (2016). Politikverdrossenheit. In L. U. Heil (Hg.), *Gesättigte Demokratie. Ein marketingorientierter Alternativbegriff zur Politikverdrossenheit* (S. 133–214). Springer VS. https://doi.org/10.1007/978-3-658-14326-8_3
- Hillgruber, C. (2016). Die Meinungsfreiheit als Grundrecht der Demokratie: Der Schutz des demokratischen Resonanzbodens in der Rechtsprechung des BVerfG. *JuristenZeitung*, 71(10), 495–501. <http://www.jstor.org/stable/24768340>
- Hinz, T., Mozer, K., & Strauß, S. (2023). *Durchführbarkeit und Qualität von rapid response research (R3HighEd). Umfragen bei Studierenden in Deutschland*. Universität Konstanz.
- Hoppe, K., Klingenberg, D., Thompson, V. E., Trautmann, F., & Vorbrugg, A. (2018). Worüber wir reden, wenn wir mit jemandem nicht reden wollen. Zum Spannungsverhältnis von Rassismuskritik und Meinungsfreiheit an der Universität. *Movements. Journal for Critical Migration and Border Regime Studies*, 4(1), 167–177. <https://doi.org/10.64081/mvmnts-4.1-2509>
- Hübner, C., Eichhorn, J., Nicke, S., & Eilers, N. (2022). *(NACH-)Wahlanalyse: Wie haben junge Deutsche 2021 bei der Bundestagswahl gewählt?* Friedrich-Ebert-Stiftung. <https://library.fes.de/pdf-files/pbud/19475.pdf>
- Infratest dimap (2024). *ARD-DeutschlandTREND August 2024. Repräsentative Studie im Auftrag der ARD*. Infratest dimap. <https://www.infratest-dimap.de/umfragen-analysen/bundesweit/ard-deutschlandtrend/2024/august/>
- Jaster, R., & Keil, G. (2021). Wen sollte man nicht an die Universität einladen? In E. Özmen (Hg.), *Wissenschaftsfreiheit im Konflikt* (S. 141–159). J. B. Metzler. https://doi.org/10.1007/978-3-662-62892-8_9

- John, S. (2021). *böll.brief. Demokratie & Gesellschaft* #25. September 2021. *Analyse der Bundestagswahl 2021*. Heinrich Böll Stiftung. https://www.boell.de/sites/default/files/2021-10/BTW21_Analyse.pdf
- Joint Committee on Human Rights. (2018). *Freedom of speech in universities – Fourth report of session 2017–19 – Report, together with formal minutes relating to the report*. HC 589 HL PAPER 111. House of Commons and House of Lords. <https://publications.parliament.uk/pa/jt201719/jtselect/jtrights/589/589.pdf>
- Kinzelbach, K., Lindberg, S. I., & Lott, L. (2024). *Academic Freedom Index 2024 Update*. FAU Erlangen-Nürnberg and V-Dem Institute. <https://doi.org/10.25593/open-fau-405>.
- Knight Foundation & Ipsos (2022). *College student views on free expression and campus speech 2022: A look at key trends in student speech views since 2016*. A Knight Foundation-Ipsos study from the Knight Free Expression Research Series. Knight Foundation & Ipsos. <https://knightfoundation.org/reports/college-student-views-on-free-expression-and-campus-speech-2022/>
- Köcher, R. (2019). *Grenzen der Freiheit. Eine Dokumentation des Beitrags von Prof. Dr. Renate Köcher in der Frankfurter Allgemeinen Zeitung Nr. 119 vom 23. Mai 2019*. Institut für Demoskopie Allensbach. https://www.ifd-allensbach.de/fileadmin/user_upload/FAZ_Mai2019_Meinungsfreiheit.pdf
- Kostner, S. (2022). Hochschulen in den 2020er-Jahren. Intellektuelle Vielfalt oder intellektuelle Lockdowns? In S. Kostner (Hg.), *Wissenschaftsfreiheit: Warum dieses Grundrecht zunehmend umkämpft ist* (S. 7–30). Nomos. <https://doi.org/10.5771/9783748928058>
- Lanius, D. (2020). Meinungsfreiheit und die kommunikative Strategie der Rechtspopulisten. In T. Schultz (Hg.), *Was darf man sagen? Meinungsfreiheit im Zeitalter des Populismus* (S. 75–112). Kohlhammer.
- Leiner, D. J. (2019). Too fast, too straight, too weird: Non-reactive indicators for meaningless data in Internet surveys. *Survey Research Methods*, 13(3), 229–248. <https://doi.org/10.18148/SRM/2019.V13I3.7403>
- Lindov, D. (2020). Teachers and politics. *European Journal of Political Economy*, 64. <https://doi.org/10.1016/j.ejpoleco.2020.101902>
- Mansour, A. (2023, 8. September). Universitäten im Spannungsfeld von Meinungsfreiheit und Aktivismus. *Jüdische Allgemeine*. <https://www.juedische-allgemeine.de/meinung/universitaeten-im-spannungsfeld-von-meinungsfreiheit-und-aktivismus/>
- Matthes, J., Knoll, J., & von Sikorski, C. (2018). The “spiral of silence” revisited: A meta-analysis on the relationship between perceptions of opinion support and political opinion expression. *Communication Research*, 45(1), 3–33. <https://doi.org/10.1177/0093650217745429>
- Meier, L. (2021). Eine soziologische Unschärferelation. Replik zum Aufsatz „Is Free Speech in Danger on University Campus? Some Preliminary Evidence from a Most Likely Case“ von Matthias Revers und Richard Traunmüller. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73, 129–135. <https://doi.org/10.1007/s11577-021-00736-0>
- Meltzer, C. E. (2017). *Medienwirkung trotz Erfahrung. Der Einfluss von direkter und medial vermittelter Erfahrung eines Ereignisses*. Springer VS. <https://doi.org/10.1007/978-3-658-15579-7>
- Menzner, J., & Traunmüller, R. (2023). Subjective freedom of speech: Why do citizens think they cannot speak freely? *Politische Vierteljahresschrift*, 64, 155–181. <https://doi.org/10.1007/s11615-022-00414-6>
- Merkley, E. (2020). Anti-intellectualism, populism, and motivated resistance to expert consensus. *Public Opinion Quarterly*, 84(1), 24–48. <https://doi.org/10.1093/poq/nfz053>

- Naughton, K. A., Eastman, N., & Perrino, N. (2017). *Speaking freely. What students think about expression at American colleges*. FIRE. Foundation for Individual Rights in Education. <https://www.fire.org/sites/default/files/2017/10/12154317/speaking-freely-2017-B-W.pdf>
- Neubaum, G., & Krämer, N. C. (2018). What do we fear? Expected sanctions for expressing minority opinions in offline and online communication. *Communication Research*, 45(2), 139–164. <https://doi.org/10.1177/0093650215623837>
- Noelle-Neumann, E. (1974). The spiral of silence. A theory of public opinion. *Journal of Communication*, 24(2), 43–51. <https://doi.org/10.1111/j.1460-2466.1974.tb00367.x>
- Noelle-Neumann, E. (1980). *Die Schweigespirale. Öffentliche Meinung – unsere soziale Haut*. Piper.
- Norris, P. (2023). Cancel culture: Myth or reality? *Political Studies*, 71(1), 145–174. <https://doi.org/10.1177/00323217211037023>
- Otto, L., & Bacherle, P. (2011). Politisches Interesse Kurzska (PIKS) – Entwicklung und Validierung. *Politische Psychologie*, 1(1), 19–35.
- Özmen, E. (Hg.) (2021). *Wissenschaftsfreiheit im Konflikt. Grundlagen, Herausforderungen und Grenzen*. J. B. Metzler.
- Petersen, T. (2020). *Forschungsfreiheit an deutschen Universitäten: Ergebnisse einer Umfrage unter Hochschullehrern. Kurzpräsentation*. Akademie der Konrad-Adenauer-Stiftung.
- Petersen, T. (2021a). *Das geistige Klima an Universitäten. Ergebnisse einer Online-Befragung von Hochschullehrern*. Institut für Demoskopie Allensbach.
- Petersen, T. (2021b). Die Mehrheit fühlt sich gegängelt. *Eine Dokumentation des Beitrags von Dr. Thomas Petersen in der Frankfurter Allgemeinen Zeitung Nr. 136 vom 16. Juni 2021*. Institut für Demoskopie Allensbach. https://www.ifd-allensbach.de/fileadmin/kurzberichte_dokumentationen/FAZ_Juni2021_Meinungsfreiheit.pdf
- Quiring, O., Jakob, N., Schemer, C., Jakobs, I., & Ziegele, M. (2020). „Das wird man doch noch sagen dürfen...“ – Wahrgenommene Sprechverbote und ihre Korrelate. In N. Jakob, O. Quiring, & M. Maurer (Hrsg.), *Traditionen und Transformationen des Öffentlichen* (S. 49–71). Springer VS. https://doi.org/10.1007/978-3-658-29321-5_3
- Revers, M., & Traunmüller, R. (2020). Is free speech in danger on university campus? Some preliminary evidence from a most likely case. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 72, 471–497. <https://doi.org/10.1007/s11577-020-00713-z>
- Rothut, S., Sacher, A.-L., Strohmeier, R., & Reinemann, C. (2023). Meinungsfreiheit in Gefahr? Wie politische Einstellungen und individuelle Erfahrungen die Wahrnehmung der Meinungsfreiheit in Deutschland prägen. *Studies in Communication and Media*, 12(1), 48–86. <https://doi.org/10.5771/2192-4007-2023-1-48>
- Sacher, A., & Reinemann, C. (2025). What is freedom of speech? How citizens define and perceive the ‘Bulwark of Liberty’. *The International Journal of Press/Politics*. <https://doi.org/10.1177/19401612251388182>
- Sachs, J. A. (2018, 16. März). The ‘campus free speech crisis’ is a myth. Here are the facts. *The Washington Post*. Abgerufen am 6. Januar 2024 von <https://www.washingtonpost.com/news/monkey-cage/wp/2018/03/16/the-campus-free-speech-crisis-is-a-myth-here-are-the-facts/>
- Sand, M., & Kunz, T. (2020). *Gewichtung in der Praxis (April 2020, Version 1.0)*. GESIS – Leibniz-Institut für Sozialwissenschaften (GESIS-Survey Guidelines). https://doi.org/10.15465/gesis-sg_030
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989X.7.2.147>

- Schieritz, M., & Scholz, A.-L. (2024, 17. Mai). Pro-Palästina-Demos: Gehen die Uni-Proteste zu weit? *Zeit Online*. Abgerufen am 21. August 2024 von <https://www.zeit.de/2024/22/pro-palaestina-demos-universitaeten-gaza-proteste-israel>
- Scott-Baumann, A., & Perfect, S. (2021). *Freedom of speech in universities: Islam, Charities and Counter-terrorism*. Routledge.
- Shell Jugendstudie. (2019). *Zusammenfassung Shell Jugendstudie*. Shell.
- Statistisches Bundesamt. (2023). *Statistischer Bericht. Statistik der Studierenden. Wintersemester 2022/2023*. Erschienen am 8. August 2023. Statistisches Bundesamt. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publicationen/Downloads-Hochschulen/statistischer-bericht-studierende-hochschulen-endg-2110410237005.xlsx?__blob=publicationFile
- Statistisches Bundesamt. (2024). *Bildung und Kultur. Studierende an Hochschulen – Fächersystematik – 2022. Übersicht 1. Fächergruppen, Studienbereiche und Studienfächer*. Stand: WS 2022/2023. Statistisches Bundesamt. https://www.destatis.de/DE/Methoden/Klassifikationen/Bildung/studenten-pruefungstatistik.pdf?__blob=publicationFile
- Stevens, S., & Haidt, J. (2018a, 4. März). The skeptics are wrong Part 1: Attitudes about free speech on campus are changing. *Heterodox Academy*. Abgerufen am 23. Januar 2024 von <https://heterodoxacademy.org/blog/skeptics-are-wrong-about-campus-speech/>
- Stevens, S., & Haidt, J. (2018b, 11. März). The skeptics are wrong part 2: Speech culture on campus is changing. *Heterodox Academy*. Abgerufen am 23. Januar 2024 von <https://heterodoxacademy.org/blog/the-skeptics-are-wrong-part-2/>
- Taddicken, M. (2013). Online-Befragung. In W. Möhring & D. Schlütz (Hrsg.), *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft* (S. 201–217). Springer VS. https://doi.org/10.1007/978-3-531-18776-1_11
- Trautmüller, R. (2022). Die »Cancel Culture«-Hypothese auf dem empirischen Prüfstand. In S. Kostner (Hg.), *Wissenschaftsfreiheit: Warum dieses Grundrecht zunehmend umkämpft ist* (S. 33–54). Nomos. <https://doi.org/10.5771/9783748928058>
- Trautmüller, R. (2023). Testing the 'campus cancel culture' hypothesis, *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4392840>
- Trautmüller, R., & Revers, M. (2021). Meinungsfreiheit an der Universität: Unschärfen und Strohmann (Antwort auf Lars Meier). *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73, 137–146. <https://doi.org/10.1007/s11577-021-00758-8>
- Unzicker, K. (2018, 17. Juli). „Das wird man ja wohl noch sagen dürfen“ – Keine Meinungsfreiheit in Deutschland? Bertelsmann Stiftung. *Blog Vielfalt leben – Gesellschaft gestalten*. Abgerufen am 3. Januar 2024 von <https://demokratie-und-zusammenhalt.de/2018/07/17/das-wird-man-ja-wohl-noch-sagen-duerfen-keine-meinungsfreiheit-in-deutschland/>
- van Buuren, S. (2018). *Flexible imputation of missing data*. Chapman & Hall/CRC. <https://stefvanbuuren.name/fimd/>
- Villa, P.-I., Trautmüller, R., & Revers, M. (2021, 12. November). Lässt sich „Cancel Culture“ empirisch belegen? Impulse für eine pluralistische Fachdebatte. *Bundeszentrale für politische Bildung*. Abgerufen am 27. Januar 2024 von <https://www.bpb.de/shop/zeitschriften/apuz/wissenschaftsfreiheit-2021/343228/laesst-sich-cancel-culture-empirisch-belegen/>
- Vogel, F. (2019). „Meinungsfreiheit“ und ihre Grenzen an der Universität. Ein Kommentar. *Navigationen – Zeitschrift für Medien- und Kulturwissenschaften*, 19(2), 33–38. <https://doi.org/10.25969/MEDIAREP/13816>
- Williams, K. D. (2007). Ostracism: The kiss of social death. *Social and Personality Psychology Compass*, 1(1), 236–247 <https://doi.org/10.1111/j.1751-9004.2007.00004.x>

- Willnat, L., Lee, W., & Detenber, B. H. (2002). Individual-level predictors of public outspokenness: A test of the spiral of silence theory in Singapore. *International Journal of Public Opinion Research*, 14(4), 391–412. <https://doi.org/10.1093/ijpor/14.4.391>
- Winterlin, F., Frischlich, L., Boberg, S., Schatto-Eckrodt, T., Reer, F., & Quandt, T. (2021). Corrective actions in the information disorder. The role of presumed media influence and hostile media perceptions for the countering of distorted user-generated content. *Political Communication*, 38(6), 773–791. <https://doi.org/10.1080/10584609.2021.1888829>
- YouGov (2018, 27. Juni). Are students really more hostile to free speech? *YouGov*. Abgerufen am 21. August 2024 von <https://yougov.com/en-gb/articles/21051-are-students-really-more-hostile-free-speech>
- ZEIT-Stiftung (2024). *Akademische Redefreiheit. Kurzbericht zu einer empirischen Studie an deutschen Hochschulen*. https://read.zeit-stiftung.com/report_akademischeredefreiheit/
- Zhou, S., & Barbaro, N. (2023). *Understanding campus expression across higher ed: Heterodox Academy's annual campus expression survey*. Heterodox Academy. <https://heterodoxacademy.org/reports/2022-campus-expression-survey-report/>

EXTENDED ABSTRACT

Free speech on campus? The impact of political congruence and fear of sanctions on perceptions of freedom of speech and willingness to speak out among German university students

Freie Rede auf dem Campus? Einflüsse von politischer Kongruenz und Sanktionsängsten auf die Meinungsfreiheitswahrnehmung und Redebereitschaft deutscher Studierender

*Luna Laufer, Franziska Selzer, Julia Stockwald, Annika Schleithoff,
Nathalie L. Neuendorf, Anna-Luisa Sacher & Carsten Reinemann*

Luna Laufer (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland.

Franziska Selzer (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland. Kontakt: franziska.selzer@ifkw.lmu.de. ORCID: <https://orcid.org/0009-0009-7572-0989>

Julia Stockwald (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland.

Annika Schleithoff, Deutsche Journalistenschule München, Hultschiner Str. 8, 81677 München, Deutschland.

Nathalie L. Neuendorf (M. A.), Universität Augsburg, Institut für Medien, Wissen und Kommunikation, Universitätsstr. 10, 86159 Augsburg, Deutschland. Kontakt: nathalie.neuendorf@uni-a.de. ORCID: <https://orcid.org/0000-0002-5938-8578>

Anna-Luisa Sacher (M. A.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland. Kontakt: anna.sacher@ifkw.lmu.de. ORCID: <https://orcid.org/0000-0002-3386-4310>

Carsten Reinemann (Prof. Dr.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstraße 67, 80538 München, Deutschland. Kontakt: carsten.reinemann@ifkw.lmu.de. ORCID: <https://orcid.org/0000-0001-9715-8434>



EXTENDED ABSTRACT

1. Theoretical framework, research questions and hypotheses

Freedom of speech is not only a fundamental human right but also an indispensable precondition for functioning liberal democracies (Garton Ash, 2016; Reinemann & Sacher, 2025). However, growing segments of the German population perceive their freedom of expression as increasingly constrained (Menzner & Traunmüller, 2023; Rothut et al., 2023). Although universities are normatively conceived as sites of free intellectual exchange and tolerance of opposing views (Scott-Baumann & Perfect, 2021; Zhou & Barbaro, 2023), campuses around the world and in Germany as well have become arenas of contested debate around “political correctness”, “wokeness, and “cancel culture”, with positions ranging from concerns about a narrowing of acceptable discourse to dismissals of such concerns as media-amplified anecdotes (Norris, 2023; Revers & Traunmüller, 2020; Traunmüller, 2023).

Previous studies of German academia have predominantly focused on perceptions and experiences of faculty members (Petersen, 2020, 2021; ZEIT-Stiftung, 2024). This has started to change only recently (Diehl et al., 2025). This study contributes to further closing this research gap by providing the first broader empirical account of free speech perceptions and participation in controversial classroom discussions among German university students across disciplines and regions. Notably, the present study is based on data collected in summer 2023, and it thus predates the events and developments following the Hamas attacks on Israel of October 7, 2023, and the war in Gaza that sparked discussions and protests at several German universities.

The study conceptualizes perceived freedom of speech as a reality perception (Rothut et al., 2023; Sacher & Reinemann, 2025) and mainly builds upon the spiral of silence theory (Noelle-Neumann, 1974). It therefore focuses on the potential effects of political congruence, i.e., the degree to which students perceive their own political orientation as aligned with that of their peers and instructors, respectively. The study hypothesizes that anticipated sanctions (e.g., expected social exclusion from peers or negative academic consequences from instructors) will be negatively associated with both free-speech perceptions and willingness to speak out in controversial classroom discussions (Neubaum & Krämer, 2018; Rothut et al., 2023).

1.1 Research questions

(FF1) How do German students perceive freedom of speech at their universities?

(FF2) How do German students in different subject areas differ in their perception of freedom of speech at their universities?

(FF3) *How frequent are classroom discussions about current controversial issues at German universities, and how often do students participate in them?*

(FF4) *How do different subject areas differ in terms of the frequency of discussions on controversial current issues and in students' participation in them?*

1.2 Hypotheses

(H1) *Higher political congruence with (a) peers and (b) instructors is associated with more positive perceptions of campus free speech.*

(H2) *Higher political congruence with (a) peers and (b) instructors is associated with greater willingness to speak out in classroom discussions.*

(H3) *Greater anticipated sanctions from (a) peers and (b) instructors are associated with more negative free-speech perceptions.*

(H4) *Greater anticipated sanctions from (a) peers and (b) instructors are associated with lower willingness to speak out.*

2. Method

Data were collected via a standardized online survey in July–August 2023 (by commercial access panel Bilendi & Respondi). After quality screening, the final weighted sample comprised $N = 897$ students from 16 German federal states and over 99 universities. Post-stratification weighting was applied to correct for the overrepresentation of humanities students and women relative to official statistics, considering gender distribution within six disciplinary clusters. To account for potential effects of weighting procedures and imputation of missing values, comprehensive robustness checks were performed with various model specifications.

Perceptions of campus free speech were captured via an eight-item scale ($\alpha = .83$, $M = 2.91$, $SD = 1.21$) adapted from Rothut et al. (2023), covering unwritten norms and taboos, the costs of honest expression, political correctness, and equal access to voice (higher values indicate more negative perceptions). Willingness to speak out was assessed via a single item on self-reported frequency of participation in controversial class discussions ($M = 2.55$, $SD = 1.00$; $N = 670$; those reporting no controversial discussions in class were excluded). Political congruence with peers and instructors was operationalized as the absolute difference between students' own left-right orientation and their estimates of peers'/instructors' political orientations. Fear of sanctions from peers (5 items, $\alpha = .75$) and instructors (6 items, $\alpha = .77$) was measured as summed dichotomous indices. Additional predictors were political orientation (1 = left, 11 = right), party preference (dummy-coded), political disaffection (composite index, $\alpha = .66$), satisfaction with democracy, political interest ($\alpha = .94$), age, and gender.

3. Results, discussion and implications

The findings challenge too alarmist narratives: Most students perceive free speech as largely intact: 71% agree (at least somewhat) that they are not excessively criticized for honest opinions, 73% agree that all have equal opportunities to be heard, and 63% note that political correctness does not suppress important discussions. However, 43% of students acknowledge “unwritten laws” about what opinions are acceptable in the courses of their main subject. No significant differences in free-speech perceptions are found across subject area clusters ($F = 1.02$, $p = .41$), suggesting that fields with more political debate are not necessarily perceived as more restrictive.

Despite universities’ normative self-image as arenas of open debate, controversial discussions are rare, and student participation is low: 23% of students report they have never encountered such discussions in class; only 12% report they occur often. Disciplinary differences are significant and substantial ($\eta^2 = .08$): Social sciences and humanities report more frequent controversy than STEM and law/economics fields. Among the 75% of students who have experienced controversial discussions, willingness to speak out is also low: 48% say they rarely or never participate; only 18% often or always engage. Participation rates do not differ significantly across disciplines, indicating that although the frequency of controversy varies, students’ willingness to speak up is similar across fields.

The regression model for free speech perceptions ($R^2 = .30$) captures political orientation as the strongest predictor ($\beta = .22$, $p < .001$): Right-leaning students perceive significantly more restrictions. Fear of sanctions from instructors is the second strongest predictor ($\beta = .19$, $p < .001$), followed by fear of sanctions from peers ($\beta = .11$, $p = .01$), confirming H3a and H3b. Political congruence with peers emerges as a significant though modest negative predictor ($\beta = -.11$, $p = .004$), partially supporting H1a. No effect is found for congruence with instructors (H1b not supported). Dissatisfaction with democracy is associated with more negative perceptions ($\beta = -.08$, $p = .02$), as in prior general population research. Among party preferences, only Green Party supporters perceive fewer restrictions ($\beta = -.14$, $p < .001$). No significant effect from supporting the right-wing populist AfD is found, in contrast to studies of the general population. Also, political disaffection is not significant in this student sample, suggesting university-specific dynamics. Finally, higher political interest ($\beta = .16$) and male gender ($\beta = -.08$) are associated with more negative free-speech perceptions.

The regression model for willingness to speak out ($R^2 = .17$) suggests that stable individual characteristics not captured here (e.g., extraversion or confidence) play a larger role in its explanation. Political disaffection appears as the strongest predictor ($\beta = -.20$, $p < .001$): Students more disaffected participate less, possibly reflecting resignation. Political interest is positively associated with participation ($\beta = .19$, $p < .001$). Neither political congruence with peers nor with instructors plays a significant role (H2a/b not supported). In contrast, fear of sanctions from peers is a robust negative predictor ($\beta = -.15$, $p = .005$; H4a supported), while the effect of instructor sanctions is not robust across model specifications (H4b not consistently supported). Counterintuitively, supporters of the AfD show signifi-

cantly higher willingness to speak out ($\beta = .14, p = .003$), as do, less robustly, SPD and Die Linke supporters – consistent with a corrective action dynamic among perceived minorities (Noelle-Neumann, 1974; Wintterlin et al., 2021). However, these effects are small and should not be overinterpreted.

Results show that the determinants of free-speech perceptions and willingness to speak out clearly are distinct phenomena: What shapes how students evaluate campus free speech is largely different from what drives their decision to speak. The partial support for spiral-of-silence mechanisms – particularly through peer-based sanction fears and, for free-speech perceptions, through political congruence with peers – extends this theoretical framework to the campus context. However, the absence of a robust political congruence effect on actual speaking behavior suggests that subjective climate perceptions matter more for evaluative judgments than for behavioral outcomes, where individual dispositions dominate. The absence of a political disaffection effect on free-speech perceptions and its negative effect on speaking behavior point to a distinct student-specific dynamic: Disaffected students may be disengaged rather than silenced.

Several limitations must be noted: Despite weighting, the non-probability sample may still contain biases, particularly with respect to politically disaffected or extreme students who may be less represented in commercial panels. The cross-sectional design precludes causal inference. The focus on classroom settings excludes broader campus life. The sanction indices are coarse, and the study does not examine the relationship between free-speech perceptions and willingness to speak out as sequentially linked constructs. Future research should (1) employ longitudinal designs, (2) incorporate personality predictors, (3) examine meso-level (institutional culture, departmental norms) and macro-level factors (media framing, political climate), and (4) illuminate the subjective meaning of free speech for students and the reasons behind their silence. Practically, the findings support calls for active cultivation of constructive debate cultures at German universities (Scott-Baumann & Perfect, 2021).

References

- Diehl, C., Revers, M., Traunmüller, R., Weidmann, N., & Wuttke, A. (2025). Students' motives for restricting academic freedom: Viewpoint discrimination and prosocial concerns. *PNAS*, 122(47). <https://doi.org/10.1073/pnas.2503804122>
- Garton Ash, T. (2016). *Free speech: Ten principles for a connected world*. Yale University Press.
- Menzner, J., & Traunmüller, R. (2023). Subjective freedom of speech: Why do citizens think they cannot speak freely? *Politische Vierteljahresschrift*, 64, 155–181. <https://doi.org/10.1007/s11615-022-00414-6>
- Neubaum, G., & Krämer, N. C. (2018). What do we fear? Expected sanctions for expressing minority opinions in offline and online communication. *Communication Research*, 45(2), 139–164.
- Noelle-Neumann, E. (1974). The spiral of silence. A theory of public opinion. *Journal of Communication*, 24(2), 43–51. <https://doi.org/10.1111/j.1460-2466.1974.tb00367.x>
- Norris, P. (2023). Cancel culture: Myth or reality? *Political Studies*, 71(1), 145–174. <https://doi.org/10.1177/00323217211037023>

- Petersen, T. (2020). *Forschungsfreiheit an deutschen Universitäten: Ergebnisse einer Umfrage unter Hochschullehrern. Kurzpräsentation* [Research freedom at German universities: Results of a survey among university lecturers]. Akademie der Konrad-Adenauer-Stiftung.
- Petersen, T. (2021). *Das geistige Klima an Universitäten. Ergebnisse einer Online-Befragung von Hochschullehrern* [The intellectual climate at universities. Results of an online survey among university lecturers]. Institut für Demoskopie Allensbach.
- Revers, M., & Traunmüller, R. (2020). Is free speech in danger on university campus? Some preliminary evidence from a most likely case. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 72, 471–497. <https://doi.org/10.1007/s11577-020-00713-z>
- Rothut, S., Sacher, A.-L., Strohmeier, R., & Reinemann, C. (2023). Meinungsfreiheit in Gefahr? Wie politische Einstellungen und individuelle Erfahrungen die Wahrnehmung der Meinungsfreiheit in Deutschland prägen [Freedom of speech under threat? How political attitudes and individual experiences shape perceptions of freedom of speech in Germany]. *Studies in Communication and Media*, 12(1), 48–86. <https://doi.org/10.5771/2192-4007-2023-1-48>
- Sacher, A., & Reinemann, C. (2025). What is freedom of speech? *The International Journal of Press/Politics*. <https://doi.org/10.1177/19401612251388182>
- Scott-Baumann, A., & Perfect, S. (2021). *Freedom of speech in universities: Islam, Charities and Counter-terrorism*. Routledge.
- Wintterlin, F., Frischlich, L., Boberg, S., Schatto-Eckrodt, T., Reer, F., & Quandt, T. (2021). Corrective actions in the information disorder. *Political Communication*, 38(6), 773–791. <https://doi.org/10.1080/10584609.2021.1888829>
- ZEIT-Stiftung (2024). *Akademische Redefreiheit. Kurzbericht zu einer empirischen Studie an deutschen Hochschulen* [Academic freedom of expression. Brief report on an empirical study at German universities]. https://read.zeit-stiftung.com/report_akademischere-defreiheit/
- Zhou, S., & Barbaro, N. (2023). *Understanding campus expression across higher ed*. Heterodox Academy.

FULL PAPER

Climate Protest—What is it good for? Linking news coverage on German climate movements and reader comments

Klimaprotest – was nützt er wirklich? Eine Untersuchung von Artikeln über deutsche Klimabewegungen und dazugehörigen Leserkomentaren

Claudia Thoms, Tobias Schrimpf & Andreas Reich

Claudia Thoms (Dr.), University of Hohenheim, Institute of Communication Science, Fruwirthstraße 46, 70599 Stuttgart, Germany. Contact: claudia.thoms@uni-hohenheim.de. ORCID: <https://orcid.org/0000-0002-6601-1170>

Tobias Schrimpf (M. A.), University of Hohenheim, Institute of Communication Science, Fruwirthstraße 49, 70599 Stuttgart, Germany. Contact: tobias.schrimpf@uni-hohenheim.de. ORCID: <https://orcid.org/0009-0001-6257-5948>

Andreas Reich (M. Sc.), University of Hohenheim, Institute of Communication Science, Fruwirthstraße 46, 70599 Stuttgart, Germany. Contact: andreas.reich@uni-hohenheim.de. ORCID: <https://orcid.org/0000-0002-2426-6490>



Climate Protest—What is it good for? Linking news coverage on German climate movements and reader comments

Klimaprotest – was nützt er wirklich? Eine Untersuchung von Artikeln über deutsche Klimabewegungen und dazugehörigen Leserkomentaren

Claudia Thoms, Tobias Schrimpf & Andreas Reich

Abstract: Media coverage plays a crucial role in shaping public perception of social movements. While attention-seeking protest tactics meet media selection criteria, they also risk distorting public perception. This study examines the German climate protest scene by comparing two key protest groups: Fridays for Future and Last Generation, the latter often criticized for its disruptive protest tactics. Building on the concept of mediatization, the protest paradigm, and agenda-setting theory, we analyze the mediated impact of different protest strategies on reader comments. Our quantitative content analysis includes 100 online news articles and 1,910 reader comments from three major German news sites (2019–2023). We find that articles about Last Generation focus more on protest actions and are less likely to prompt discussions about the climate issue among commentators. They are also more negative than those about Fridays for Future. However, journalistic evaluations are not directly reflected in the corresponding reader comments. These results suggest that non-normative and disruptive protest strategies could harm the discourse on climate change. Our findings are relevant for social movements and interest groups seeking to raise awareness of socially relevant issues such as climate protection and climate change. Additionally, this study contributes to understanding the dynamics between social movements, the media, and the public within the political communication system.

Keywords: Climate movement, protest strategy, mediatization, agenda-setting, content analysis

Zusammenfassung: Medienberichterstattung prägt die öffentliche Wahrnehmung von sozialen Bewegungen maßgeblich. Spektakuläre Protestaktionen erfüllen zwar die medialen Selektionskriterien, bergen aber auch die Gefahr, dass sie die öffentliche Wahrnehmung verzerren. Diese Studie untersucht die deutsche Klimabewegung anhand zweier Gruppen: Fridays for Future und Letzte Generation, wobei letztere oft für ihre disruptiven Protesttaktiken kritisiert wird. Aufbauend auf dem Konzept der Mediatisierung, dem Protest Paradigm und der Agenda-Setting-Theorie analysieren wir den medienvermittelten Effekt verschiedener Proteststrategien auf Leserkommentare. Unsere quantitative Inhaltsanalyse umfasst 100 Online-Nachrichtenartikel und 1.910 Leserkommentare von drei großen deutschen Nachrichtenseiten (2019–2023). Artikel über die Letzte Generation konzentrieren sich dabei mehr auf deren Protestaktionen und führen in den Leserkomentaren selte-

ner zu Diskussionen über das Klimathema. Sie sind auch negativer als Artikel, die Fridays for Future betreffen. Dagegen spiegeln sich journalistische Bewertungen nicht direkt in den dazugehörigen Leserkomentaren wider. Diese Ergebnisse legen nahe, dass nicht-normative und disruptive Proteststrategien dem Diskurs über den Klimawandel schaden könnten. Unsere Ergebnisse sind relevant für soziale Bewegungen und Interessengruppen, die versuchen, das Bewusstsein für gesellschaftlich relevante Themen wie Klimaschutz und Klimawandel zu schärfen. Darüber hinaus trägt diese Studie zum Verständnis der Dynamik zwischen sozialen Bewegungen, den Medien und der Öffentlichkeit innerhalb des politischen Kommunikationssystems bei.

Schlagwörter: Klimabewegung, Proteststrategie, Mediatisierung, Agenda Setting, Inhaltsanalyse

1. Introduction

Social movements, like other non-media actors, rely on media attention to disseminate their messages to the public, achieve their goals, and contribute to democratic opinion-forming (Strömbäck & Esser, 2017). Attention-seeking protest action can capture media attention. However, the more disruptive such protest actions are, the more they risk diverting focus from the actual cause, potentially leading to the delegitimization and demonization of the movement (McLeod, 2007, pp. 186–187). Therefore, a key challenge for social movements consists of finding protest forms that meet journalistic criteria of newsworthiness without detracting from their core message (Goldenbaum & Thompson, 2020, p. 198). As such, they need to adapt to the media *strategically*. This is crucial, as the public's perception, evaluation, and support of social movements, their activities, and messages depend decisively on the media coverage they receive (e.g., Brown & Mourão, 2021, p. 591), ultimately influencing the movements' social impact (e.g., Brehm & Gruhl, 2024, p. 5). Given this, studying the interplay between social movements' strategic choices and their influence on media and public discourse becomes especially relevant when pressing issues like climate change are concerned.

In recent years, two influential actors have shaped Germany's climate protest scene: *Fridays for Future* (FFF) and *Letzte Generation*¹ (Last Generation, LG). Although they advocate for similar issues, their protest strategies² differ. FFF relies on peaceful and cheerful protests and school strikes, whereas LG engages in civil disobedience and more disruptive protests, such as blocking roads and vandalizing public monuments. This has led to comparisons that reveal remarkable differences in how the protest groups are treated. For example, Kumkar (2022, p. 1) and Rucht (2023, pp. 17–25) observed that media reports frequently allege that LG's protest strategy shifts the focus away from climate issues, causing defi-

1 Now renamed to *Neue Generation* (New Generation).

2 We understand *protest strategies* as a summary of protest actions that follow a certain logic, while *protest tactics* represent specific forms of action, like demonstrations or street blockades (similarly: Doherty, 2013). Thus, we use the terms *protest tactics* and *protest actions* interchangeably. The main difference between the two is their applicability to specific protest events, which can also be referred to as protest actions, but not protest tactics.

ance rather than solidarity towards the movement and discouraging public support for the climate cause. Recent studies analyzing media portrayals of both protest groups showed that such criticisms are common. In general, journalists tend to pay more attention to LG than to FFF, but they are also more critical of LG (Dablander et al., 2025, p. 11; Meyer, Farjam et al., 2025, pp. 16–18). Furthermore, the described allegations about the negative impact of such disruptive protest strategies on public perceptions of and public discussions about the protest groups have been confirmed, at least partially, as LG faces social resistance (Rucht, 2023, p. 19). In contrast, despite receiving limited political support beyond parties aligned with their views (Berker & Pollex, 2023, pp. 768–769), FFF has gained widespread societal acceptance (Rucht, 2023, p. 17). Moreover, an examination of online discussions concerning the two protest groups indicated that LG's disruptive protests also resulted in heightened polarization of the discourse (e.g., Meyer, Pröschel et al., 2025, p. 14).

While the influence of media portrayals on public discourse is generally recognized in such research, studies establishing a direct link between the two are rare and often designed in a way that makes it difficult to trace a direct connection between *specific* media portrayals of the two protest groups and public reactions to them (e.g., Brehm & Gruhl, 2024, p. 5; Heide, 2025, pp. 985–986). We address this gap by taking advantage of interactive online media to study how readers react to different media portrayals of the two protest groups and their protests. Our case study employs a quantitative content analysis of 100 news articles and 1,910 reader comments to trace the connections between protest strategy, media coverage, and reader perceptions about the two climate movements FFF and LG in Germany. Based on this analysis, we want to answer two questions. First, how does the media craft narratives about the movements, considering their different protest strategies? As mentioned, several studies have been published on this question lately. However, this analysis is a necessary condition for addressing our second, more crucial question regarding the research gap described above: How do these narratives impact public discussions about the movements, their goals, protest actions, and the broader climate topic? In the following, the concept of mediatization, the protest paradigm, and agenda-setting theory serve as a basis for understanding the connection between protest strategy, media coverage, and public opinion.

2. Literature review

2.1 Mediatization and the communication strategies of social movements

Social movements are “sustained campaign[s] of claim making, using repeated performances that advertise the claim, based on organizations, networks, traditions, and solidarities that sustain these activities” (Tilly & Tarrow, 2015, p. 11). They aim to evoke social or political change by rallying the masses to their goal (Rucht, 2012, p. 3). This is because public support can lend weight to their concerns (Rasmussen et al., 2018, pp. 156–157). To achieve this, public attention is an important prerequisite (Hunt & Gruszczynski, 2021, p. 1027).

Although social media allows social movements to bypass journalistic gatekeepers (Hunt & Gruszczynski, 2021, p. 1035), traditional mass media (hereafter referred to as media in contrast to social media) remain crucial for attracting, amplifying, and sustaining attention to issues that shape both political (Langer & Gruber, 2021) and public (Hunt & Gruszczynski, 2021) agendas. A promising strategy for managing media attention involves either meeting their expectations or intentionally disregarding their needs and wants (Scheu, 2019, pp. 4–5). The concept of mediatization, as understood by the institutionalist tradition, describes this strategic adaptation to the selection, presentation, and quality criteria that characterize journalistic information dissemination, also known as media logic (Haßler, 2017, pp. 25–26, 82; Strömbäck & Van Aelst, 2013, p. 345). Importantly, such *strategic* adaptation to journalistic criteria of newsworthiness involves *purposeful* adaptation, meaning that it should serve to achieve the movement's goals (Strömbäck & Esser, 2017, p. 74). Avoiding media attention by deliberately refraining from communicating in a particularly media-like manner can be useful in this sense. This strategy, also known as *defensive mediatization* (Scheu, 2019, pp. 4–5), has been discussed as a possible choice for social movements in previous research (McCurdy, 2012, pp. 249–250). However, according to Strehler-Schaaf (2025, pp. 280–281), this choice is more due to negative experiences or a lack of resources than to a disregard for the importance of public communication in general. As she points out, it is possible to identify protest movements that are skeptical or cautious in their approach to the media, i.e., movements that pursue a more defensive strategy. Yet, these movements are in the minority compared to those that are more offensive. Furthermore, in her analysis, environmental movements fall into the category of media-challenging protest movements that notably adapt to media logic (Strehler-Schaaf, 2025, pp. 245–254). As a result, we focus on strategies that fall under the category of *offensive mediatization*. Since social movements rely heavily on media coverage to reach the public (Jennings & Saunders, 2019, p. 2305), their best option to increase their chances of publicity lies in adapting media logic as intensively as possible (Scheu, 2019, p. 4). As news values and routines have a significant impact on the coverage of social movements, it is crucial for these movements to follow the “rules of the ‘media game’” (Vliegert-hart & Walgrave, 2012, p. 390). This allows them to overcome their systematic disadvantages compared to institutionalized political actors with greater public access. Indeed, features of protest events, such as greater numbers of protesters and increased levels of violence between police and protesters, may lead to heightened media coverage, particularly in the short term (Jennings & Saunders, 2019, p. 2306). These features are related to news factors like relevance or reach in the former case, and damage, controversy, conflict, or aggression in the latter, which are generally influential in journalistic (and audience) news selection (Eilders, 2006, p. 11).

Thus, attention-seeking protest actions can serve as an offensive mediatization strategy and attract media attention. More extreme protest actions can increase visibility and attention (Feinberg et al., 2020, p. 1089). However, the more disruptive protest actions are, the more they also risk diverting focus from the actual cause (e.g., Smith et al., 2001, pp. 1415–1417) to “violence and disruption, de-

emphasis of the protesters' voices, and invocation of negative comments from bystanders" (Lee, 2014, p. 2737), potentially leading to the delegitimization, demonization, and marginalization of the protest movement. The protest paradigm, as "a routinized template for creating protest stories that has been naturalized through the process of journalistic socialization" (McLeod, 2007, p. 186), highlights this tension. Coverage following this pattern has remained on a high level over time (Gruber, 2023, p. 428). As such, it can lead the public to view the protests as unnecessary, as the specific way they are presented makes it difficult to understand the protesters' perspective. Additionally, Feinberg et al. (2020, p. 1106) demonstrated that protest tactics perceived as harmful to others, disruptive, or both can diminish public support for both the movement and its positions by reducing social identification with the movement. Ultimately, this may also harm the potential for mobilization.

Consequently, while there may not be a universally right way to protest in terms of achieving different goals simultaneously (e.g., garnering high levels of media attention and favorable framing), the strategic decisions made between one protest tactic and another can still make a difference (Shuman et al., 2024). McLeod's (2007, p. 194) recommendations to activists suggest alternatives for adapting to media logic while reducing the risk of falling into the protest paradigm. For example, mobilizing larger crowds (e.g., through social media) or enlisting prominent figures as supporters can be beneficial in terms of news factors such as reach or prominence, as well as elite persons (e.g., Nardini et al., 2021, p. 128). It can make the protest activities interesting for the media. Additionally, compared to more disruptive forms of protest, this kind of mobilization can effectively signal to the media that the protesting group is not just a group of deviants, making it harder to portray and marginalize them as such (McLeod, 2007, p. 194; Wouters & Van Camp, 2017, p. 451). According to Zlobina and Gonzalez Vazquez (2018, p. 245), people are more willing to participate in protests the more they perceive the protest tactic as both morally right (i.e., legitimate) and useful (i.e., effective). When comparing different protest tactics, the authors show that demonstrations as a "prototypically normative form of protest" (p. 245) are considered the most legitimate and have a higher likelihood of participation. Shuman et al.'s (2024, p. 260) review essentially confirms this, arguing that normative nonviolent protests are better for mobilization and that non-normative or violent protests tend to be more effective at advancing policy change. Thus, protest groups must consider that following the social rules of the game (e.g., through the choice of announced demonstrations) and engaging in rule-breaking behavior, which may include civil disobedience, are perceived differently and prompt different outcomes. However, normativeness is not the only dimension along which forms of protest can differ (for different categorizations, see Andrews & Caren, 2010; Camaj, 2014; Caren et al., 2020; Evans, 2016; McCluskey, 2008; Zlobina & Gonzalez Vazquez, 2018). Another common distinction is between disruptive, attention-grabbing forms of protest, such as blocking roads, and more authoritative strategies, which social movements can also consider using instead. Authoritative strategies are aimed at gaining recognition and authority, allowing activists to position themselves as experts or key representatives of a constituency (Caren

et al., 2020, p. 446). These include conventional media relations tools, such as publishing press releases or contacting reporters (Andrews & Caren, 2010, pp. 846, 854; Caren et al., 2020, p. 446). As Evans (2016, p. 43) pointed out, the use of this kind of tactics can demonstrate professionalism. This professionalism attracts journalists and helps the movement to be recognized as “a legitimate spokesman for a particular interest” (Mirowsky & Ross, 1981, pp. 178, 187). One way to reinforce this is by employing specialized personnel with the necessary expertise (Andrews & Caren, 2010, p. 854) or by providing some kind of leader (e.g., Nepstad & Bob, 2006, p. 18). Not only does this meet the media’s demands for personalization, but it also ideally provides them with a reliable communication partner. Finally, clear, concise, and coherent communication is crucial for both the media and the audience to understand the movement’s objectives (McLeod, 2007, p. 194). This is consistent with the fact that comprehensibility is a key quality criterion in journalism, to which non-media actors can also adapt and from which the public can benefit (Thoms, 2023, pp. 58–63).

However, even if a social movement is perfectly attuned to the media, there is no guarantee of success in the battle for media attention. External factors such as the political environment (Caren et al., 2020, p. 447) and “being in tune with the media agenda” (Wouters & Lefevere, 2023, p. 628) are crucial for a social movement, even though they cannot be directly controlled. Adding to this, the political orientation of the media outlet can also influence the occurrence and strength of the protest paradigm, with conservative outlets being more prone to it than liberal ones (Lee, 2014, p. 2737; Meyer, Farjam et al., 2025, pp. 16–17). Nevertheless, the strategic decisions of a social movement may still lead to differences in the journalistic reactions they evoke. As Meyer, Farjam et al. (2025, pp. 13, 16–17) demonstrated in their comparison of media coverage of FFF and LG, the disruptiveness of protest tactics is another important factor influencing the prevalence of negative framing, at least partly overriding the effect of the media’s political orientation.

2.2 The agenda-setting function of media’s coverage of social movements

As previously stated, reaching out to the media is not an end in itself, but a means to draw political and public attention to a movement’s concerns and to build support. The media’s agenda-setting function plays a crucial role in this context (McCombs & Shaw, 1972). From the perspective of communication science, the effect of the media on the public agenda (as compared to that of political actors) is of particular interest and still relevant despite the development of a high-choice media environment (Djerf-Pierre & Shehata, 2017, p. 751).³

The agenda-setting theory posits that the quantity, prominence, and tone of media coverage can increase public awareness of certain issues (traditional agenda-setting effects) or aspects of those issues (attribute agenda-setting effects; see

3 The effect of the media on political actors is more commonly studied in political science. According to Van Aelst et al. (2014, pp. 202–203), despite some general similarities, there are significant differences between political and public agenda-setting.

Kiousis, 2004, pp. 74–76; McCombs, 2005, pp. 544–547). In this way, “the media not only can be successful in telling us *what to think about*, they also can be successful in telling us *how to think about it*” (McCombs, 2005, p. 546). As Kiousis (2004, p. 76) argues, attributes can be both substantive (cognitive) and affective (emotional); the media can thus tell us not only how to *think* but also how to *feel* about an issue. Studying social movements and their coverage, both agenda-setting perspectives are relevant. First, the coverage of such protests can influence not only how the public perceives the movement itself but also the issue it is addressing. Second, regarding the protest paradigm, the emphasis that media coverage places on different aspects of protests may differentially influence the public.

Researchers have studied agenda-setting effects using different methods across different cultural contexts. Examining the relations between media content and survey data, Smidt’s (2012, p. 88) study revealed that media coverage of demonstrations related to gun control and Obama’s health care reform had an impact on public attention to these issues. Wallace et al. (2014, pp. 444–445) examined the relationship between media use and political alienation among respondents, focusing on perceptions of the 2006 protests for immigrant rights in the USA. While the study lacks content analytical data, it is suspected that how the protests were portrayed in parts of the media may have led to the negative attitudes of Latino respondents. Furthermore, Camaj (2014, p. 650) showed that attribute agenda-setting has an impact on respondents’ political participation in both institutional (e.g., voting) and non-institutional forms (e.g., demonstrations) of participation. This result is also relevant for social movements and supports the findings of Zlobina and Gonzalez Vazquez (2018), as well as Feinberg et al. (2020). It underscores the importance of a positive portrayal and, in a second step, a positive evaluation of a protest to increase the likelihood that people will join or support it. Based on an extensive literature review examining different outcomes of climate protests, Thomas-Walters et al. (2025, pp. 2–3) concluded that there is strong evidence for effects of climate protests on media coverage and public opinion, but only weak evidence for agenda-setting effects. However, Brehm and Gruhl (2024, p. 5) emphasized the mediating role of media coverage of climate protests in increasing public concerns about climate change. Nevertheless, they highlighted the importance of considering different media frames, which their study design does not address.

The rise of the Internet and the digitalization of the media landscape have created opportunities for alternative measures of the public agenda that can be helpful in this instance. Reader comments can be used to track direct public responses to *specific* news items. This highlights the potential benefits of interactive online media for studying agenda-setting effects by demonstrating how media content influences readers’ discussions (e.g., Danner et al., 2022; Roberts et al., 2002; with results contrary to general agenda-setting research findings, see Koltsova & Nagornyy, 2019). Concerning protests, Kleut and Milojevic (2021, p. 95) showed that the frames used in media coverage of the Serbian protest *Against Dictatorship* were widely echoed in reader comments. More closely related to climate change, Walter et al. (2018, pp. 213–214) found that the article’s position on the

topic seemed to have less influence on reader comments than expected. Just as Lörcher and Taddicken (2017, p. 14) write: “public discourse is provoked by journalists, but follows different patterns of perception”.

Although reader comments represent only a specific part of the public (Kalogeropoulos et al., 2017, pp. 7–8), their analysis offers an important advantage: The ability to observe direct reactions to (many) different journalistic presentations of social movements and their protests. Other methods require significantly more effort to fulfill this goal, which is why they often focus on aggregated effects of media use (i.e., through surveys) or are necessarily restricted to analyzing the effects of a few stimuli (i.e., through experiments).

3. Climate protest: The case of FFF and LG

For our case study, we compare two climate movements that are active in Germany. FFF is an international climate movement that has been active in Germany since 2018, focusing on school strikes (Buzogány & Scherhauser, 2022). Since then, it has expanded its repertoire to include other protest tactics, like demonstrations and educational events (Fridays for Future, 2024). Its main objective is to ensure that the target of limiting global warming to 1.5° C, as set out in the *Paris Climate Agreement*, is not exceeded.

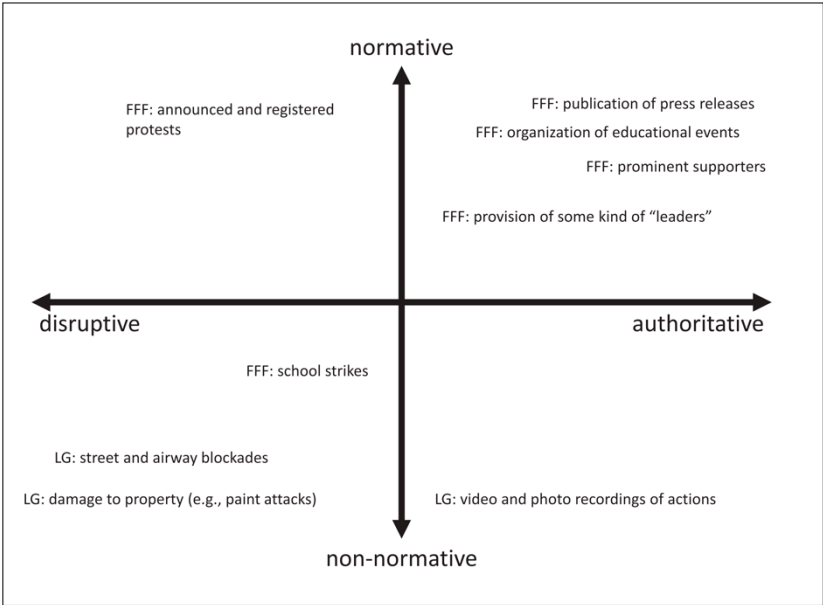
LG was founded in 2021 and has since gained public attention for its protest tactics of civil disobedience, such as blocking roads while being glued to the street or painting galleries and monuments like the Brandenburg Gate (SWR Aktuell, 2023). Their main goal is to reduce CO₂ emissions (Letzte Generation, 2024). Calling for more radical protest forms than demonstrations to shift public attention to the climate issue, the group has since been investigated by the German *Bundesamt für Verfassungsschutz* (Federal Office for the Protection of the Constitution). However, it has not yet been classified as extremist (Küstner, 2023).

We decided to analyze FFF and LG because they advocate for similar issues but pursue different protest strategies. While both distance themselves from violent forms of protest, they use different tactics to gain media attention. Those protest tactics may be the “driving force behind news coverage” (Boyle et al., 2012, p. 137). Figure 1 distinguishes FFF and LG according to their most prominent protest tactics.⁴ Based on the reviewed literature, we posit that the activities of the two protest groups can be situated along two dimensions: Normative versus non-normative and authoritative versus disruptive. LG’s most prominent protest actions, such as blockades and paint attacks, lean heavily toward the non-normative and disruptive end of the spectrum. FFF, on the other hand, focuses more on normative forms, which also include a strong element of authoritative tactics. Both can be considered representatives of different types of offensive mediatization, as described above. However, FFF manages to conform to media logic while avoiding protest actions that starkly trigger the protest paradigm. LG is often confronted with the accusation that their protest tactics impede understanding of

4 The placement of the activities of the two movements along the two dimensions is qualitative.

the climate issue and foster frustration, harming the cause in the end (Molter, 2023). Brouwer et al. (2022) argue that the public’s defensive reaction might be a normal response to “rebel-induced threat” (p. 6). As they explain, this can be followed up by understanding and a change of perspective under the right conditions (e.g., focusing on the audience’s self-efficacy in contributing to a more climate-friendly society). It is unclear, however, whether FFF or LG meet these conditions. Bugden (2020, p. 9) further argues that social movements may use normative (demonstrations) and non-normative (civil disobedience) tactics in a complementary way. His study found no evidence of so-called backfire effects regarding both forms of protest.

Figure 1. Classification of the most prominent tactics of the respective movement



4. Research questions and hypotheses

We assume that social movements typically adopt an offensive mediatization strategy because they depend on media attention to achieve their goals. Disruptive protest tactics may be disadvantageous due to the underlying mechanisms of the protest paradigm as well as the public perception and evaluation of related protest actions. We focus on climate protests as a case study and compare two influential actors that have shaped Germany’s climate protest scene in recent years: FFF and LG. The differing protest strategies of the two protest groups have led to comparisons, revealing notable distinctions. Therefore, our first research question is:

RQ1: What are the differences between FFF coverage and LG coverage?

As mentioned above, recent studies analyzing media portrayals of both protest groups revealed the prevalence of the protest paradigm. Due to this, LG's use of non-normative and disruptive protest tactics has been shown to have a negative impact on their media coverage compared to FFF. This results in a more negative portrayal of the movement and a stronger emphasis on their protest actions rather than their concerns. As the first step in our broader analysis, we expect to also see this in our data. Specifically, we expect the following:

Hypothesis 1.1: Articles about LG show a more negative evaluation of the movement than articles about FFF.

Hypothesis 1.2: Articles about LG provide more negative evaluations of the protest actions than articles about FFF.

Hypothesis 1.3: Articles about LG are more likely to solely focus on protest actions instead of also discussing the movement's goals and/or the topic of climate change compared to articles about FFF.

As noted above, attracting media attention is essential, but it is only the first step in building public support for a movement and its causes. Various protest strategies may lead to distinct media coverage, providing their readers with different angles to focus on in subsequent discussions. Therefore, our second research question is:

RQ2: How does the reporting style affect commentator discussion?

In this context, we are interested in the extent to which media ratings influence reader perception. Additionally, we want to know if the accusation frequently made in public discourse, that LG's protest has a detrimental effect on engagement with the climate issue, has any merit. This is in line with the assumptions of the protest paradigm, meaning that LG's disruptive protest tactics lead media coverage to focus on certain (negative) aspects of protests, thereby influencing follow-up communication through the mechanisms of agenda-setting. Specifically, we expect the following:

Hypothesis 2.1: Evaluations of (a) the movements, (b) their goals, and (c) their protest actions in articles are positively associated with evaluations of these protest attributes in reader comments.

Hypothesis 2.2: Articles that solely focus on protest actions instead of also discussing the movement's goals and/or the issue of climate change decrease the likelihood that commentators will talk about the issue of climate change.

Hypothesis 2.3: Articles that solely focus on protest actions instead of also discussing the movement's goals and/or the issue of climate change lead to a more negative evaluation of the issue of climate change in reader comments.

5. Method

5.1 Data collection and coding procedure

To test our hypotheses, we conducted a quantitative content analysis of news articles and corresponding reader comments about FFF and LG. For this, we scraped online articles from three German news sites: *TAZ*, *Zeit*, and *Welt*. These are digital offshoots of leading national publications that cover a range of editorial lines. Research has shown that editorial lines can influence the occurrence of the protest paradigm, which is why we chose to include a broader ideological spectrum with *TAZ* being left-wing, *Zeit* left-liberal, and *Welt* conservative (euro-topics, n.d.). All three news sites provide comment sections for their articles. Searching for articles using the keywords “Fridays for Future” and “Letzte Generation” yielded a total of 3,485 articles, also including opinion pieces. We excluded articles that lacked text elements or reader comments, leaving 2,477 articles. A total of 122,291 reader comments were extracted for these articles, using only the comments and no replies to comments. Finally, we omitted duplicates and articles with fewer than 20 reader comments to ensure a minimum level of engagement with the corresponding articles. In addition, we only included articles published between January 1, 2019, and September 30, 2023, for FFF and between January 1, 2021, and September 30, 2023, for LG. This resulted in a total of 1,379 articles (966 for FFF; 413 for LG) and 108,286 reader comments. There was an overrepresentation of articles from *Zeit* and articles related to FFF.

To keep the sample size manageable for manual coding, we drew a random sample. This was done in two consecutive waves. In the first step, a stratified disproportional random sampling method was used to counteract the aforementioned overrepresentation. We drew 102 articles and 20 reader comments each. Care was taken to ensure equal representation of groups and news sites at the article level. Some articles, however, were deemed irrelevant during coding, requiring a redraw. Through oversampling, the target of 102 articles was ultimately achieved. Each article was coded exclusively for the group it was drawn from during sampling. Two articles, which were extracted both for FFF and LG, were excluded from the sample.⁵ Overall, a slight imbalance in distribution between the protest groups and news sites remained in the final 100 articles (*TAZ*: $n_{FFF} = 19$, $n_{LG} = 17$; *Zeit*: $n_{FFF} = 14$, $n_{LG} = 18$; *Welt*: $n_{FFF} = 15$, $n_{LG} = 17$). A total of 1,910 reader comments were coded as relevant. These form our basis at the comment level (*TAZ*: $n_{FFF} = 368$, $n_{LG} = 331$; *Zeit*: $n_{FFF} = 253$, $n_{LG} = 328$; *Welt*: $n_{FFF} = 293$, $n_{LG} = 337$).

Three trained coders coded the material based on a codebook developed for the study.⁶ The coding categories used for the article level were also applied to the

5 Previously, articles that were drawn for both groups (a total of 55 from 1,379 articles) were assigned to one group. Unfortunately, this was done through the data wrangling process to the detriment of LG. Two of these articles, now assigned to FFF, ended up in the sample, which is why we excluded them. This only led to a small imbalance between the two groups ($n_{FFF} = 48$, $n_{LG} = 52$).

6 The German language codebook is accessible upon request.

comment level, but with necessary adjustments made to account for the differences in text types. For example, articles allow for more detailed explanations, while reader comments are often briefer, requiring more reading between the lines. In addition, readers may express opinions in their comments that diverge from the prevailing societal and scientific consensus, such as questioning the veracity of climate change. This is highly unlikely for journalists writing for the type of media analyzed. The coders' reliability was tested based on the coding of a sample of 30 articles and 120 reader comments. Four reliability coefficients were considered: *Holsti*, *Lotus*, *Standardized Lotus* (chance-correcting version of *Lotus*; Fretwurst, 2013), and *Krippendorff's Alpha*. The strictest coefficient, Krippendorff's Alpha, shows critical values below .67 for several variables at both coding levels (e.g., thematization of goals and climate change at the article level and evaluation of climate change at the reader comment level). However, most variables showed Holsti values of at least .70 at the article level, except for thematization of goals, thematization of climate change, and evaluation of goals. According to *Lotus*, these still exhibit a high level of agreement with the most frequently coded value. Thus, while better reliability values would be preferable, we consider the coding quality acceptable. Details on the variables relevant to our hypotheses, their reliability, and their distribution in the final sample are presented below and in Table A1 in the Appendix.

5.2 Measures

Focus of the article. The focus of the article was operationalized through a binary coding of the presence of three elements: Movements' protest actions, movement's goals, and the thematization of the climate change issue. The coding of protest actions depended on whether the text mentioned any protest tactics (e.g., school strikes, demonstrations, damage to property) used by the movements to draw attention to their goals. To code the thematization of the movement's goals, the text had to clearly state that it referred to the movement's demands. Accusations against political leaders, companies, or parts of society could also be interpreted as goals, as they are implicitly linked to a demand. The article was coded as referencing climate change if it discussed the topic in general terms, including its consequences and measures taken to address it. Simply mentioning the terms "climate change" or "climate protection" was insufficient. Based on this, a new variable was created to indicate whether an article focused exclusively on protest actions (1; $n = 25$) or deviated from this focus by discussing goals or the climate issue alongside the protest actions or by omitting any reference to a protest action (0; $n = 75$).

Evaluation of the movements, their goals, and their protest actions in the articles and the reader comments. Our focus is on implicit and explicit evaluations. We consider not only the author's evaluations, but also the positions or views of quoted individuals (e.g., when a politician is quoted criticizing or praising a particular goal). In other words, we consider all evaluations in the article when assessing the evaluated object. Thus, we focused on the tone of the texts and not on the valence of the events or protest actions described in them. We categorized

evaluations as negative (-1), ambivalent or neutral (0), and positive (1). In exceptional cases, a comment could be classified as unclear if it was imprecise or ambiguous (-99). Typically, the evaluation of a movement is related to the evaluation of its protest actions and goals. However, it can also be more general, underscoring the importance of separate coding.

Thematization of climate change in the reader comment. This was coded similarly to the coding on the article level (see the description under *Focus of the article*).

Evaluation of climate change in the reader comment. Commentators may provide varying evaluations regarding climate change. To assess the commentary's stance on climate change, we considered both direct and subtle assessments, including those based on specific word choices (e.g., the term "climate change" may be used neutrally, while "climate crisis" may convey a more negative connotation). The evaluation of climate change was deemed negative or irrelevant (-1) if it was denied or considered less urgent, with climate change being downplayed explicitly. A positive or relevant (1) rating was given if the crisis nature of the issue or the urgency of action was emphasized. A neutral (0) presentation lacked any explicit or implicit evaluation. Again, in exceptional cases, a comment could be classified as unclear if it was imprecise or ambiguous (-99).

5.3 Analysis

Our data has a nested structure in which reader comments are linked to articles. RQ1 and its associated hypotheses only pertain to the article level. Therefore, we tested them using regular (linear and logistic) regression models. In contrast, RQ2 and its associated hypotheses involve both levels. Because reader comments are a reaction to an article and its characteristics, reader comments on one article will be more similar to each other than to comments on another article. This lack of independence in our observations affects the results of statistical analyses, leading us to use multilevel regression models, specifically random intercept models with articles treated as random effects. Indeed, the *intraclass correlations* (ICC) estimated on the intercept-only models demonstrated the validity of this decision, as they indicated that between 27–30% of the variance of the dependent variables at the comment level could be explained by characteristics at the article level.

Dependent variables for RQ1 included the evaluation of protest movements and their actions, as well as the focus of the article on protest actions. Regarding RQ2, the analysis included the evaluation of protest movements, their goals, and protest actions, as well as the thematization and evaluation of climate change in reader comments. Article characteristics like the protest group and the news source for RQ1 and RQ2, as well as the evaluation of the protest movement, their goals, and protest actions, combined with the focus of the article on protest actions for RQ2, were included in the models as fixed effects to examine their relationship with comment characteristics. The two protest groups and the news sites were included in all models to account for any systematic differences in the results caused by these factors. We coded FFF and TAZ as reference groups and LG, Welt, and Zeit as treatment groups. Therefore, the coefficients associated with the

treatment indicators quantify the deviance of FFF from LG, TAZ from Welt, and TAZ from Zeit, respectively. For all models, we observed a significantly better fit compared to the null models. More complex models (e.g., using random slopes) were not considered, since random intercept models already reflected the theoretical considerations of the hypotheses.

6. Results

6.1 Differences between FFF coverage and LG coverage

We find that articles about LG tend to take a more negative stance on the movement compared to FFF, even when controlling for the news site from which the article originated (H1.1, see Table 1). As expected, we observe that Welt and Zeit take a more negative stance than the left-leaning TAZ, with the conservative Welt having a more pronounced negative effect on the evaluation of the protest movement than the left-liberal Zeit. The same applies to the evaluation of the protest actions of the respective movement (H1.2): LG's protest actions are rated more negatively overall than those of FFF. This could also be due to the fact that various protest actions have more positive or negative connotations overall. We therefore investigated how often the different protest actions appeared in articles about the two movements to determine which protest actions are most strongly associated with the respective protest movement: Articles covering protest actions of FFF ($n = 34$) mainly focus on strikes (53%), demonstrations (15%), or other actions (29%), while articles about protest actions of LG ($n = 51$) are dominated by stories about street blockades (69%), property damage (10%), and other actions (14%). LG's protest actions, which are much more disruptive, could therefore damage the overall impression of the movement.

Furthermore, we find that, compared to FFF, articles about LG more often focus entirely on the movement's protest actions rather than other topics such as its goals (38% vs. 10%, H1.3). This suggests that the protest strategy chosen by LG, which relies heavily on attention-generating protest actions such as street blockades or vandalism, is not as effective in placing actual demands or goals in the media coverage.

To summarize, we find evidence for all our first three hypotheses addressing the differences between FFF coverage and LG coverage (RQ1).

Table 1. Linear and logistic regression models predicting article level evaluations and focus

	Article level		
	Evaluation of movements (H1.1)	Evaluation of protest actions (H1.2)	Sole focus on protest actions (H1.3)
	<i>b</i> (SE)	<i>b</i> (SE)	OR (SE)
(Intercept)	0.76 (0.11) ***	0.51 (0.14) ***	0.11 (0.06) ***
Group: LG	-0.50 (0.12) ***	-0.53 (0.14) ***	5.32 (2.94) **
News site: Welt	-0.71 (0.14) ***	-0.55 (0.17) **	1.08 (0.65)
News site: Zeit	-0.42 (0.14) **	-0.24 (0.18)	1.22 (0.73)
<i>N</i> _{articles}	96	85	100
<i>R</i> ² / <i>R</i> ² adjusted	0.34 / 0.31	0.23 / 0.21	0.11

Notes. For H1.3, *R*² corresponds to *Tjur's R*² for logistic regressions. Values in bold are significant. * *p* < .05 ** *p* < .01 *** *p* < .001.

6.2 Media effects on commentator discussion

Turning to our second research question, where we analyze how the reporting style might affect commentator discussions, we find that, contrary to our initial hypothesis, holding protest groups and news sites constant, journalists' evaluations of (a) movements, (b) their goals, and (c) their protest actions are not related to those of commentators (see Table 2).⁷ Thus, we find no support for H2.1. Instead, the observation of Walter et al. (2018) about the lack of influence of individual news stories on the way readers comment can be confirmed in our setting. It is probable that their preexisting opinions are more influential than the evaluations they encounter when reading a specific article. We continue to see that the news sites have a (negative) impact regarding Welt and Zeit, although the latter shows no significant influence on the evaluation of the goals and protest actions.

7 Because missing values due to the coding logic resulted in different sample sizes, we did not include all article level evaluations in each model.

Table 2. Multilevel regression models predicting reader comment level evaluations

	Reader comment level		
	Evaluation of movements (H2.1a)	Evaluation of goals (H2.1b)	Evaluation of protest actions (H2.1c)
	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)
Fixed part			
(Intercept)	0.34 (0.07) ***	0.55 (0.19) **	0.13 (0.10)
Article level evaluation of ...			
... movements	-0.09 (0.05)		
... goals		0.05 (0.15)	
... protest actions			-0.07 (0.07)
Group: LG	-0.25 (0.07) ***	-0.22 (0.16)	-0.29 (0.10) **
News site: Welt	-1.01 (0.08) ***	-1.12 (0.20) ***	-0.72 (0.11) ***
News site: Zeit	-0.27 (0.08) ***	-0.19 (0.21)	-0.23 (0.12)
Random part			
τ_{00}	0.03	0.07	0.08
σ^2	0.48	0.61	0.48
$N_{articles}$	88	43	75
$N_{comments}$	942	148	594
Marginal R^2 / Conditional R^2	0.26 / 0.30	0.26 / 0.33	0.15 / 0.28

Notes. τ_{00} = between-group random intercept variance; σ^2 = residual variance. Values in bold are significant. * $p < .05$ ** $p < .01$ *** $p < .001$.

Given the results associated with H2.1, the results of H2.3 are similarly unsurprising. The thematic focus of the article is of little consequence in determining how readers evaluate the issue of climate change in their comments (see Table 3). It can be concluded that the concern that the (often negative or critical) portrayal of protest actions in the media could have a *direct* negative impact on attitudes towards the issue of climate change in general is unfounded. H2.3 is not supported. Nevertheless, we continue to see a negative effect if the article and subsequent discussions are conducted on Welt, but not on Zeit.

However, the results on the likelihood of thematizing the issue of climate change in the reader comments if there is a focus on protest actions in the article (H2.2) certainly fuel the concern that the protest actions may have a cannibalizing effect (see Table 3): If the focus of the coverage is on the protest actions, climate change is less likely to be discussed in the reader comments.

Table 3. Multilevel regression models predicting reader comment level thematization and evaluation of climate change

	Reader comment level	
	Thematization of climate change (H2.2)	Evaluation of climate change (H2.3)
	OR (SE)	<i>b</i> (SE)
Fixed part		
(Intercept)	0.46 (0.08) ***	0.61 (0.07) ***
Article level solely focuses on protest actions	0.33 (0.10) ***	-0.14 (0.12)
Group: LG	0.55 (0.14) *	0.22 (0.08) **
News site: Welt		-0.81 (0.10) ***
News site: Zeit		-0.14 (0.09)
Random part		
τ_{00}	1.02	0.03
σ^2	3.29	0.45
$N_{articles}$	100	87
$N_{comments}$	1,908	473
Marginal R^2 / Conditional R^2	0.09 / 0.30	0.20 / 0.26

Notes. τ_{00} = between-group random intercept variance; σ^2 = residual variance. Values in bold are significant. For H2.2, the full model did not converge, which is why the control variable for news sites was excluded here. A logistic regression with all control variables, but without taking the multiple levels into account, yielded similar results concerning the influence of the article's focus (see Appendix Table A2). * $p < .05$ ** $p < .01$ *** $p < .001$.

7. Discussion

The purpose of our study was to examine the impact of two different protest strategies on media coverage and readers' reactions. From the perspective of social movements, it is important to understand whether their communication activities hinder or support their goals and causes. As public perceptions of climate protest could influence how issues like climate change are perceived and whether related policies are supported, this is also a matter of societal importance. We compared evaluations and thematic priorities in news coverage across different editorial lines and reader discussions about two climate movements that are active in Germany and differ starkly in their protest strategies: FFF and LG. While both follow an offensive mediatization strategy, LG is leaning more towards non-normative and disruptive protest tactics. Research shows that these tactics are more prone to prompt the so-called protest paradigm, with possible repercussions for public perceptions of social movements. The design of our case study allows us to draw first conclusions on the effects of media content on follow-up communication, since we rely on the most direct public reaction to media coverage,

namely reader comments. This approach provides an innovative perspective on public perception of FFF and LG. However, our findings also align with existing research.

We find that news articles report more negatively about LG compared to FFF (Table 1). News articles about LG focus more strongly on their disruptive and partly illegal protest actions and less often on the goals of the movement. FFF, with its chosen protest strategy of focusing less on disruptive protest actions, seems to place its goals more often in the reporting. These initial findings are in line with current research (Dablander et al., 2025; Meyer, Farjam et al., 2025) and lend weight to critics' claims that LG's protest strategy is failing to raise awareness of the climate issue. In contrast, a more balanced protest strategy that relies more heavily on normative and authoritative tactics appears to be advantageous for successfully placing the climate topic in news articles.

Furthermore, journalistic evaluations of the protest groups are not directly reflected in the corresponding reader comments. The overall more negative evaluation at the comment level compared to the article level indicates a more critical view of the readers than of the journalists (see Appendix Table A1). Such opinions, however, are likely to be the consequence of previous media consumption (e.g., Brown & Mourão, 2021, p. 587) and interpersonal communication. The results of H1.1 and H1.2 indicate that readers of all three news sites are exposed to a greater number of negative articles about LG compared to FFF, which, as we mentioned, comparable research supports. This is likely driving our results.

Similarly, the media's focus on protest actions does not appear to directly induce climate skepticism in readers or to attract those with relevant predispositions. As previously stated, while the individual article may be less relevant to the formation of a stable opinion, the cumulative effect of media coverage can be. The results pertaining to the effect of the focus on protest actions on readers' engagement with the climate change issue are pertinent in this context. When the coverage centers solely on protest actions, commentators are less likely to discuss climate change compared to instances where the movement's goals or climate change in general are mentioned. This suggests that coverage focused on protest actions diverts attention away from the issue. Different protest strategies and the associated forms of protest can indeed lead to different follow-up communication due to variations in their media coverage. This ties in well with agenda-setting research.

The findings of our first research question align with the protest paradigm, which has been demonstrated to be prevalent across diverse journalistic cultures (Boyle et al., 2012, p. 139), suggesting some generalizability of our findings. In the long term, a strategic emphasis on non-normative and disruptive protest actions by social movements may, in fact, have a detrimental impact on the climate movement and the climate cause. Although these can be successful in generating attention (Feinberg et al., 2020, p. 1089), they can displace a more general discussion of the issue and divert public attention from the causes and goals behind the protest (Smith et al., 2001, pp. 1415–1417), which is in line with the protest paradigm. This can hinder public understanding and support. The findings of our

second research question indicate that this is possible, as the readers' direct reaction to media coverage shows.

There are limitations to our study, offering interesting opportunities for follow-up studies. First, although our results are significant, a larger sample size at the article level would be beneficial to make more reliable statements. Recent developments in the automated analysis of texts using large language models seem promising in this respect (e.g., Kroon et al., 2024). Second, concerning intercoder reliability of some categories, our results should be interpreted with some caution. Nevertheless, the results are consistent with similar research, which supports the validity of our findings. Third, despite its advantages, studying readers' comments means focusing on a part of the public that is not representative of the population, even when considering media with different political orientations. As demonstrated in an international comparative study by Kalogeropoulos et al. (2017, pp. 7–8), the segment of the audience that comments on news is not only small but also demographically, attitudinally, and behaviorally different from the non-commenting majority. More general statements about the public agenda, therefore, depend on additional, ideally fine-grained surveys that include non-commentators. Fourth, we did not consider the influences commentators can exert on each other and on the perception of news. The perceived interestingness of a discussion influences the extent to which other commentators participate in the discussion (Hoppe et al., 2019, p. 209). Negative comments, in turn, can have a negative impact on the credibility of the article and the perceived importance of the topic (Waddell, 2018, p. 3080). This could affect our question about the likelihood that commentators will talk about the issue of climate change. Investigating these effects would require a more in-depth analysis of the dynamics in the comment sections and the relationship of these dynamics to the ratings of the article or its topic over time.

While it is reasonable to assume that people form stable opinions based on past media coverage, it would be beneficial to track changes in coverage and reader comments related to protests over time. This would be worthwhile to develop a better understanding of opinion formation processes. Combined with the inclusion of additional actors and/or climate change coverage beyond thematically related protests, it could also paint a more complete picture of what the public is learning about climate change and climate activism in general. Finally, strategic changes of movements call for comparisons within a movement. Given the influence of preconceptions not only on journalists but also on the public, the question is whether the change in protest strategy announced by LG (Marg, 2024, pp. 14–16), for example, will actually pay off or be good for absolutely nothing.

Acknowledgements

We would like to thank Tarek Bahiani, Sophie Mann, and Sabine Sonnemann for their assistance with the content analysis. We also thank the anonymous reviewers for their constructive and detailed feedback, which greatly helped to improve this paper.

Generative AI declaration

We acknowledge the use of AI-based tools during the preparation of this manuscript. ChatGPT (publicly available version at the time of data collection, October/November 2023) was used to support the coding of the web scraper. DeepL software was used to translate and refine the English text of this manuscript. The authors carefully reviewed and proofread all code and text to ensure accuracy and correctness.

Literature

- Andrews, K. T., & Caren, N. (2010). Making the news: Movement organizations, media attention, and the public agenda. *American Sociological Review*, 75(6), 841–866. <https://doi.org/10.1177/0003122410386689>
- Berker, L. E., & Pollex, J. (2023). Explaining differences in party reactions to the Fridays for Future-movement – A qualitative comparative analysis (QCA) of parties in three European countries. *Environmental Politics*, 32(5), 755–792. <https://doi.org/10.1080/09644016.2022.2127536>
- Boyle, M. P., McLeod, D. M., & Armstrong, C. L. (2012). Adherence to the protest paradigm: The influence of protest goals and tactics on news coverage in U.S. and international newspapers. *The International Journal of Press/Politics*, 17(2), 127–144. <https://doi.org/10.1177/19401612111433837>
- Brehm, J., & Gruhl, H. (2024). Increase in concerns about climate change following climate strikes and civil disobedience in Germany. *Nature Communications*, 15(1). <https://doi.org/10.1038/s41467-024-46477-4>
- Brouwer, C., Bolderdijk, J.-W., Cornelissen, G., & Kurz, T. (2022). Communication strategies for moral rebels: How to talk about change in order to inspire self-efficacy in others. *WIREs Climate Change*, 13(5). <https://doi.org/10.1002/wcc.781>
- Brown, D. K., & Mourão, R. R. (2021). Protest coverage matters: How media framing and visual communication affects support for black civil rights protests. *Mass Communication and Society*, 24(4), 576–596. <https://doi.org/10.1080/15205436.2021.1884724>
- Bugden, D. (2020). Does climate protest work? Partisanship, protest, and sentiment pools. *Socius*, 6. <https://doi.org/10.1177/2378023120925949>
- Buzogány, A., & Scherhauser, P. (2022). Framing different energy futures? Comparing Fridays for Future and Extinction Rebellion in Germany. *Futures*, 137. <https://doi.org/10.1016/j.futures.2022.102904>
- Camaj, L. (2014). The consequences of attribute agenda-setting effects for political trust, participation, and protest behavior. *Journal of Broadcasting & Electronic Media*, 58(4), 634–654. <https://doi.org/10.1080/08838151.2014.966363>
- Caren, N., Andrews, K. T., & Lu, T. (2020). Contemporary social movements in a hybrid media environment. *Annual Review of Sociology*, 46(1), 443–465. <https://doi.org/10.1146/annurev-soc-121919-054627>
- Dablander, F., Wimmer, S., & Haslbeck, J. (2025). Media coverage of climate activist groups in Germany. *Climatic Change*, 178(8), 1–46. <https://doi.org/10.1007/s10584-025-03959-8>
- Danner, H., Hagerer, G., Pan, Y., & Groh, G. (2022). The news media and its audience: Agenda setting on organic food in the United States and Germany. *Journal of Cleaner Production*, 354. <https://doi.org/10.1016/j.jclepro.2022.131503>

- Djerf-Pierre, M., & Shehata, A. (2017). Still an agenda setter: Traditional news media and public opinion during the transition from low to high choice media environments. *Journal of Communication*, 67(5), 733–757. <https://doi.org/10.1111/jcom.12327>
- Doherty, B. (2013). Tactics. In D. A. Snow, D. della Porta, B. Klandermans, & D. McAdam (Eds.), *The Wiley-Blackwell Encyclopedia of Social and Political Movements*. Wiley. <https://doi.org/10.1002/9780470674871.wbespm210>
- Eilders, C. (2006). News factors and news decisions. Theoretical and methodological advances in Germany. *Communications*, 31(1), 5–24. <https://doi.org/10.1515/COMMUN.2006.002>
- eurotopics. (n.d.). *Media*. eurotopics. <https://www.eurotopics.net/en/142186/media>
- Evans, E. M. (2016). Bearing witness: How controversial organizations get the media coverage they want. *Social Movement Studies*, 15(1), 41–59. <https://doi.org/10.1080/14742837.2015.1060158>
- Feinberg, M., Willer, R., & Kovacheff, C. (2020). The activist's dilemma: Extreme protest actions reduce popular support for social movements. *Journal of Personality and Social Psychology*, 119(5), 1086–1111. <https://doi.org/10.1037/pspi0000230>
- Fretwurst, B. (2013). *Intercoderreliabilität und Expertenvalidität bei Inhaltsanalysen: Erläuterungen zur Berechnung des Reliabilitätskoeffizienten Lotus mit SPSS* [Intercoder reliability and expert validity in content analyses: Explanations for calculating the Lotus reliability coefficient with SPSS]. Institut für angewandte Kommunikationsforschung. <https://iakom.ch/documents/6/LotusManualGer.pdf>
- Fridays for Future. (2024). *Aktionen* [Actions]. <https://fridaysforfuture.de/aktionen/>
- Goldenbaum, M., & Thompson, C. S. (2020). Fridays for Future im Spiegel der Medienöffentlichkeit [Fridays for Future as reflected in the media]. In S. Haunss & M. Sommer (Eds.), *Fridays for Future – Die Jugend gegen den Klimawandel* (pp. 181–204). transcript. <https://doi.org/10.1515/9783839453476-009>
- Gruber, J. B. (2023). Troublemakers in the streets? A framing analysis of newspaper coverage of protests in the UK 1992–2017. *The International Journal of Press/Politics*, 28(2), 414–433. <https://doi.org/10.1177/19401612221102058>
- Haßler, J. (2017). *Mediatisierung der Klimapolitik* [Mediatization of climate policy]. Springer VS. <https://doi.org/10.1007/978-3-658-15668-8>
- Heide, J. (2025). Polarizing social figures? Climate activists in German media and popular discourse. *European Societies*, 27(5), 959–991. https://doi.org/10.1162/euso_a_00030
- Hoppe, I., Lörcher, I., & Kießling, B. (2019). Und die Welt schaut (wieder) hin? Agenda-Setting-Effekte klimabezogener Ereignisse in zwei Online-Öffentlichkeitsarenen [And the world is watching (again)? Agenda-setting effects of climate-related events in two online public arenas]. In I. Neverla, M. Taddicken, I. Lörcher, & I. Hoppe (Eds.), *Klimawandel im Kopf: Studien zur Wirkung, Aneignung und Online-Kommunikation* (pp. 203–228). Springer Fachmedien. https://doi.org/10.1007/978-3-658-22145-4_8
- Hunt, K., & Gruszczynski, M. (2021). The influence of new and traditional media coverage on public attention to social movements: The case of the Dakota Access Pipeline protests. *Information, Communication & Society*, 24(7), 1024–1040. <https://doi.org/10.1080/1369118X.2019.1670228>
- Jennings, W., & Saunders, C. (2019). Street demonstrations and the media agenda: An analysis of the dynamics of protest agenda setting. *Comparative Political Studies*, 52(13–14), 2283–2313. <https://doi.org/10.1177/0010414019830736>
- Kalogeropoulos, A., Negrodo, S., Picone, I., & Nielsen, R. K. (2017). Who shares and comments on news?: A cross-national comparative analysis of online and social media participation. *Social Media + Society*, 3(4), 1–12. <https://doi.org/10.1177/2056305117735754>

- Kiousis, S. (2004). Explicating media salience: A factor analysis of New York Times issue coverage during the 2000 U.S. presidential election. *Journal of Communication*, 54(1), 71–87. <https://doi.org/10.1111/j.1460-2466.2004.tb02614.x>
- Kleut, J., & Milojevic, A. (2021). Framing protest in online news and readers' comments: The case of Serbian protest "Against Dictatorship". *International Journal of Communication*, 15, 82–102.
- Koltsova, O., & Nagornyy, O. (2019). Redefining media agendas: Topic problematization in online reader comments. *Media and Communication*, 7(3), 145–156. <https://doi.org/10.17645/mac.v7i3.1894>
- Kroon, A., Welbers, K., Trilling, D., & van Atteveldt, W. (2024). Advancing automated content analysis for a new era of media effects research: The key role of transfer learning. *Communication Methods and Measures*, 18(2), 142–162. <https://doi.org/10.1080/19312458.2023.2261372>
- Kumkar, N. C. (2022, November 16). Die Radikalisierung der Radikalisierungsbehauptung: Zum Diskurs über die Letzte Generation [The radicalization of the radicalization claim: On the discourse about the Last Generation]. *Soziopolis: Gesellschaft beobachten*. <https://www.sozipolis.de/die-radikalisierung-der-radikalisierungsbehauptung.html>
- Küstner, K. (2023, March 15). "Letzte Generation" bisher "nicht extremistisch" ["Last Generation" so far "not extremist"]. *tagesschau*. Retrieved from <https://www.tagesschau.de/inland/haldenwang-verfassungsschutz-letzte-generation-101.html>
- Langer, A. I., & Gruber, J. B. (2021). Political agenda setting in the hybrid media system: Why legacy media still matter a great deal. *The International Journal of Press/Politics*, 26(2), 313–340. <https://doi.org/10.1177/1940161220925023>
- Lee, F. L. F. (2014). Triggering the protest paradigm: Examining factors affecting news coverage of protests. *International Journal of Communication*, 8, 2725–2746.
- Letzte Generation. (2024). *Gesellschaftsrat + Fragen & Antworten* [Citizens' council + questions & answers]. <https://letztegeneration.org/gesellschaftsrat/>
- Lörcher, I., & Taddicken, M. (2017). Discussing climate change online. Topics and perceptions in online climate change communication in different online public arenas. *Journal of Science Communication*, 16(2). <https://doi.org/10.22323/2.16020203>
- Marg, S. (2024). *Kurzstudie: Die Letzte Generation. Entwicklung, Merkmale und Einordnung* (Stand Februar 2024) [Brief study: The Last Generation. Development, characteristics and classification (as of February 2024)]. Institut für Demokratieforschung – Bundesfachstelle Linke Militanz. <https://www.linke-militanz.de/data/akten/2024/03/kurzstudie-die-letzte-generation-marg-2024.pdf>
- McCluskey, M. R. (2008). Activist group attributes and their influences on news portrayal. *Journalism & Mass Communication Quarterly*, 85(4), 769–784. <https://doi.org/10.1177/107769900808500404>
- McCombs, M. (2005). A look at agenda-setting: Past, present and future. *Journalism Studies*, 6(4), 543–557. <https://doi.org/10.1080/14616700500250438>
- McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176–187. <https://doi.org/10.1086/267990>
- McCurdy, P. (2012). Social movements, protest and mainstream media. *Sociology Compass*, 6(3), 244–255. <https://doi.org/10.1111/j.1751-9020.2011.00448.x>
- McLeod, D. M. (2007). News coverage and social protest: How the media's protest paradigm exacerbates social conflict. *Journal of Dispute Resolution*, 1, 185–194.
- Meyer, H., Farjam, M., Rauxloh, H., & Brüggemann, M. (2025). From disruptive protests to disrupted news frames: Comparing German news on climate protests. *Journalism*. Advance online publication. <https://doi.org/10.1177/14648849251372805>

- Meyer, H., Pröschel, L., & Brüggemann, M. (2025). From disruptive protests to disrupted networks? Analyzing levels of polarization in the German Twitter/X debates on “Fridays for Future” and “Letzte Generation”. *Social Media + Society*, 11(2), 1–16. <https://doi.org/10.1177/20563051251337400>
- Mirowsky, J., & Ross, C. E. (1981). Protest group success: The impact of group characteristics, social control, and context. *Sociological Focus*, 14(3), 177–192. <https://doi.org/10.1080/00380237.1981.10570394>
- Molter, N. (2023, May 29). Acht von zehn Deutschen: „Letzte Generation“ tut Klimaschutz keinen Gefallen [Eight out of ten Germans: “Last Generation” is not doing climate protection any favors]. *Augsburger Allgemeine*. Retrieved from <https://www.augsburger-allgemeine.de/special/bayern-monitor/umfrage-acht-von-zehn-deutschen-letzte-generation-tut-klimaschutz-keinen-gefallen-id66635936.html>
- Nardini, G., Rank-Christman, T., Bublitz, M. G., Cross, S. N. N., & Peracchio, L. A. (2021). Together we rise: How social movements succeed. *Journal of Consumer Psychology*, 31(1), 112–145. <https://doi.org/10.1002/jcpy.1201>
- Nepstad, S. E., & Bob, C. (2006). When do leaders matter? Hypotheses on leadership dynamics in social movements. *Mobilization: An International Quarterly*, 11(1), 1–22. <https://doi.org/10.17813/mai.11.1.013313600164m727>
- Rasmussen, A., Mäder, L. K., & Reher, S. (2018). With a little help from the people? The role of public opinion in advocacy success. *Comparative Political Studies*, 51(2), 139–164. <https://doi.org/10.1177/0010414017695334>
- Roberts, M., Wanta, W., & Dzwo, T.-H. (D.). (2002). Agenda setting and issue salience online. *Communication Research*, 29(4), 452–465. <https://doi.org/10.1177/0093650202029004004>
- Rucht, D. (2012). Massen mobilisieren [Mobilize the masses]. *APuZ*, 62(25–26), 3–9.
- Rucht, D. (2023). *Die Letzte Generation: Beschreibung und Kritik* [The Last Generation: Description and review] (ipb working series 1/2023). ipb. https://protestinstitut.eu/wp-content/uploads/2023/04/WP_1.2023.pdf
- Scheu, A. M. (2019). Between offensive and defensive mediatization. An exploration of mediatization strategies of German science-policy stakeholders. *Journal of Science Communication*, 18(3). <https://doi.org/10.22323/2.18030208>
- Shuman, E., Goldenberg, A., Saguy, T., Halperin, E., & van Zomeren, M. (2024). When are social protests effective? *Trends in Cognitive Sciences*, 28(3), 252–263. <https://doi.org/10.1016/j.tics.2023.10.003>
- Smidt, C. D. (2012). Not all news is the same: Protests, presidents, and the mass public agenda. *Public Opinion Quarterly*, 76(1), 72–94. <https://doi.org/10.1093/poq/nfr019>
- Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From protest to agenda building: Description bias in media coverage of protest events in Washington, D.C. *Social Forces*, 79(4), 1397–1423. <https://doi.org/10.1353/sof.2001.0053>
- Strehler-Schaaf, M. (2025). *Proteste für Zeitungen, Fernsehen und Insta? Wie sich Protestbewegungen (nicht) an wahrgenommene Medienlogiken anpassen* [Protests for newspapers, television, and Instagram? How protest movements (do not) adapt to perceived media logic]. Springer VS. <https://doi.org/10.1007/978-3-658-49959-4>
- Strömbäck, J., & Esser, F. (2017). Political public relations and mediatization: The strategies of news management. In P. Van Aelst & S. Walgrave (Eds.), *How Political Actors Use the Media: A Functional Analysis of the Media's Role in Politics* (pp. 63–83). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-60249-3_4
- Strömbäck, J., & Van Aelst, P. (2013). Why political parties adapt to the media: Exploring the fourth dimension of mediatization. *International Communication Gazette*, 75(4), 341–358. <https://doi.org/10.1177/1748048513482266>

- SWR Aktuell. (2023, July 17). Aktiver Widerstand. Finanzierung, Ziele, Hintergründe: Das ist die "Letzte Generation" [Active resistance. Financing, goals, background: This is the "Last Generation"]. *SWR Aktuell*. Retrieved from <https://web.archive.org/web/20250319025721/https://www.swr.de/swraktuell/letzte-generation-klimaaktivisten-fragen-antworten-100.html>
- Thomas-Walters, L., Scheuch, E. G., Ong, A., & Goldberg, M. H. (2025). The impacts of climate activism. *Current Opinion in Behavioral Sciences*, 63. <https://doi.org/10.1016/j.cobeha.2025.101498>
- Thoms, C. (2023). *Im Sinne der Medien – Textverständlichkeit im Nachrichtenauswahlkontext* [Serving media preferences – text comprehensibility in the news selection context]. Springer VS. <https://doi.org/10.1007/978-3-658-40007-1>
- Tilly, C., & Tarrow, S. (2015). *Contentious politics*. Oxford University Press.
- Van Aelst, P., Thesen, G., Walgrave, S., & Vliegenthart, R. (2014). Mediatization and political agenda-setting: Changing issue priorities? In F. Esser & J. Strömbäck (Eds.), *Mediatization of Politics: Understanding the Transformation of Western Democracies* (pp. 200–220). Palgrave Macmillan. https://doi.org/10.1057/9781137275844_11
- Vliegenthart, R., & Walgrave, S. (2012). The interdependency of mass media and social movements. In H. A. Semetko & M. Scammell (Eds.), *The SAGE Handbook of Political Communication* (pp. 387–397). SAGE. <https://doi.org/10.4135/9781446201015.n31>
- Waddell, T. F. (2018). What does the crowd think? How online comments and popularity metrics affect news credibility and issue importance. *New Media & Society*, 20(8), 3068–3083. <https://doi.org/10.1177/1461444817742905>
- Wallace, S. J., Zepeda-Millán, C., & Jones-Correa, M. (2014). Spatial and temporal proximity: Examining the effects of protests on political attitudes. *American Journal of Political Science*, 58(2), 433–448. <https://doi.org/10.1111/ajps.12060>
- Walter, S., Brüggemann, M., & Engesser, S. (2018). Echo chambers of denial: Explaining user comments on climate change. *Environmental Communication*, 12(2), 204–217. <https://doi.org/10.1080/17524032.2017.1394893>
- Wouters, R., & Lefevere, J. (2023). Making their mark? How protest sparks, surfs, and sustains media issue attention. *Political Communication*, 40(5), 615–632. <https://doi.org/10.1080/10584609.2023.2188499>
- Wouters, R., & Van Camp, K. (2017). Less than expected? How media cover demonstration turnout. *The International Journal of Press/Politics*, 22(4), 450–470. <https://doi.org/10.1177/1940161217720773>
- Zlobina, A., & Gonzalez Vazquez, A. (2018). What is the right way to protest? On the process of justification of protest, and its relationship to the propensity to participate in different types of protest. *Social Movement Studies*, 17(2), 234–250. <https://doi.org/10.1080/14742837.2017.1393408>

Appendix

Table A1. Reliability of the variables and distribution in the final sample

	Holsti	Lotus	Std. Lotus	Krippen- dorff's Alpha	% or M (SD)	N
<i>Article level</i>						
Focus: Thematization of ...						
... goals	.64	.82	.64	.28	58%	100
... protest actions	.97	.99	.97	.84	85%	100
... climate change	.56	.78	.56	.11	52%	100
Evaluation of ...						
... movements	.81	.90	.87	.84	0.14 (0.69)	96
... goals	.57	.78	.70	.54	0.47 (0.57)	58
... protest actions	.79	.89	.85	.77	-0.06 (0.73)	85
<i>Reader comment level</i>						
Thematization of climate change	.82	.91	.82	.61	25%	1,908
Evaluation of ...						
... movements	.71	.84	.78	.65	-0.28 (0.83)	946
... goals	.84	.92	.89	.65	0.07 (0.92)	174
... protest actions	.77	.88	.83	.70	-0.38 (0.82)	608
... climate change	.72	.85	.80	.45	0.44 (0.77)	473

Notes. Reliability coefficients were calculated based on 24 articles and 113 comments that were deemed relevant. The reliability of the relevance coding was high on both levels (article level: Holsti: .93, Lotus: .97, Standardized Lotus: .93, Krippendorff's Alpha: .75; reader comment level: Holsti: .98, Lotus: .99, Standardized Lotus: .98, Krippendorff's Alpha: .79). Missing values resulting from dependencies in the coding procedure were treated as a decision and recoded with -99 to be included in the comparison of coders for all coefficients except Krippendorff's Alpha, since it can handle missing values. Additionally, evaluations were treated as ordinal for Krippendorff's Alpha. Percentage (%), mean (M), and standard deviation (SD) refer to the distribution of the original variables in the final sample (100 articles and 1,910 comments). N refers to the total number of valid cases at the article and reader comment level.

Table A2. Alternative regression model predicting reader comment level thematization of climate change

	Reader comment level thematization of climate change (H2.2)
	OR (SE)
(Intercept)	0.67 (0.06) ***
Article level solely focuses on protest actions	0.39 (0.06) ***
Group: LG	0.56 (0.06) ***
News site: Welt	0.51 (0.07) ***
News site: Zeit	0.93 (0.12)
$N_{comments}$	1,908
Tjur's R^2	0.06

Note. * $p < .05$ ** $p < .01$ *** $p < .001$

FULL PAPER

Influential film producers and their networks: The formation of creative teams and gender inequality

Einflussreiche Filmproduzenten und ihre Netzwerke: Die Zusammensetzung von Filmteams und Geschlechterungleichheit

Elizabeth Prommer, Deb Verhoeven & Skadi Loist

Elizabeth Prommer (Prof. Dr.), University of Rostock, Institute for Media Research, August-Bebel-Str. 28, 18055 Rostock, Germany. Contact: elizabeth.prommer@uni-rostock.de. ORCID: <https://orcid.org/0000-0001-5238-4884>

Deb Verhoeven (Prof. Dr.), University of Alberta, Canada 150 Laureate, Women's and Gender Studies, 3–51 Assiniboia Hall, T6G 2E7 Edmonton, Alberta, Canada. Contact: debver@ualberta.ca. ORCID: <https://orcid.org/0000-0003-3680-3561>

Skadi Loist (Prof. Dr.), Norwegian University of Science and Technology (NTNU), Department of Art and Media Studies, Høgskoleringen 1, 7491 Trondheim, Norway. Contact: skadi.loist@ntnu.no. ORCID: <https://orcid.org/0000-0002-8299-4103>



Influential film producers and their networks: The formation of creative teams and gender inequality

Einflussreiche Filmproduzenten und ihre Netzwerke: Die Zusammensetzung von Filmteams und Geschlechterungleichheit

Elizabeth Prommer, Deb Verhoeven & Skadi Loist

Abstract: Gender inequality in screen industries across the world is well researched at both global and national scales. Evidence-based evaluation of inequality in these industries typically occurs at the level of annualised, aggregated statistical analysis in the form of summative data describing the number or proportion of women and gender minorities present in the screen sectors. Instead, this research study proceeds from the basis that gender inequality stems from unequal *social* relationships. We use Social Network Analysis (SNA) to evaluate the specific nature of collaboration networks in the German film industry over an eight-year period. We then extended the SNA research with career biographies, focussing our attention on the collaborative behaviours of the industry's most influential beneficiaries: male film producers. Whereas previous research concluded there is no "pipeline" problem in the film industry because women graduate from film schools at statistically equitable levels, our findings reveal that the professional relationships formed between men at film school persist throughout their careers and operate to obstruct women's opportunities.

Keywords: Film industry, gender, Social Network Analysis, gender equity, film school

Zusammenfassung: Geschlechterungleichheit in der Film- und Fernsehbranche ist sowohl auf globaler als auch auf nationaler Ebene gut erforscht. Eine evidenzbasierte Bewertung dieser Ungleichheit erfolgt in der Regel über jährlich aggregierte Statistiken, die in Form summarischer Daten die Anzahl oder den Anteil von Frauen und geschlechtlichen Minderheiten in der Branche erfassen. Diese Studie hingegen geht davon aus, dass Geschlechterungleichheit aus ungleichen *sozialen* Beziehungen resultiert. Wir wenden Soziale Netzwerkanalyse (SNA) an, um die spezifische Struktur von Kollaborationsnetzwerken in der deutschen Filmindustrie über einen Zeitraum von acht Jahren zu untersuchen. Ergänzend zur SNA beziehen wir biografische Informationen ein und richten unser Augenmerk insbesondere auf die Kooperationsmuster der einflussreichsten Akteure der Branche: männliche Filmproduzenten. Während frühere Studien zu dem Schluss kamen, dass es kein „Nachwuchsproblem“ gebe, da Frauen Filmhochschulen in statistisch gleichem Maße wie Männer abschließen, zeigen unsere Ergebnisse, dass die während des Filmstudiums geknüpften professionellen Beziehungen unter Männern über ihre gesamte Laufbahn hinweg bestehen bleiben und als strukturelles Hindernis für die Karrieren von Frauen wirken.

Schlagwörter: Filmindustrie, Geschlecht, Soziale Netzwerkanalyse, Geschlechtergerechtigkeit, Filmhochschule

1. Introduction and background

Gender inequality in the film industry is well researched. Empirical data show that women produce less audiovisual content as directors, producers, and writers worldwide and are seen less on screen (Lauzen, 2021; Loist et al., 2024; Prommer & Loist, 2020; Ross & Padovani, 2017; Simone, 2024; Smith, Choueiti, Pieper et al., 2020; Smith, Choueiti, Yao et al., 2020). The same data suggest that this is also not the result of a “pipeline” problem, as women make up at least 50 percent of graduates from film schools (Aylett et al., 2016). In an overview of gender inequality in Australia, the Australian screen agency confirmed that, “the cause of the imbalance does not appear to lie in education or training; rather, it lies in what happens after education ends.” (Screen Australia, 2015) But what happens after film school that leads to the underrepresentation of women? Previous research has shown that film schools are a key entry point into the industry, where professional connections are established (Banks, 2019; Jones et al., 2026). Our specific interest lies in the question, whether film schools play an important role in the formation of gender-biased creative film teams and, consequently, whether exclusion mechanisms begin prior to entry into the industry.

Gender equality and diversity in the film industry matter for several reasons. In addition to basic issues of workplace equity, several studies suggest that gender inequality in production teams leads to unequal visibility of diversity on screen (Liddy, 2020; Linke & Prommer, 2021). Laura Mulvey (1975) demonstrated how male directors adopt the “male gaze”, in which the story, camera angle and reception of a film are composed with a male audience in mind. In order to see diverse worlds on screen, it is necessary to have a diversity of content producers, as bell hooks argues (hooks, 2015).

A recent summary report on gender inequality in the creative industries, commissioned by UNESCO, notes:

Diverse cultural expressions that promote gender equality can transform our individual and collective perceptions, banish long-held and damaging stereotypes and amplify the voices and stories of the invisible and/or long-silenced. (Conor, 2021, p. 5)

To date, the international film industries have been impervious to equity-driven interventions, and the longitudinal data show little improvement over time. Industry and academic reports, often commissioned by national funding agencies, illustrate the snail’s pace in change for a wide range of countries (Flicker & Vogelmann, 2018). Overall, studies have demonstrated that the argument that there are “no good films by women” or “there is no female talent” does not hold (Aylett et al., 2016; Hochfeld et al., 2017; Loist, 2018; Loist & Prommer, 2019). Instead, what the data suggests is that severe structural barriers exist for women during their careers, which restrict their ability to enter the industry and/or main-

tain a career, and which must be further considered and addressed by researchers and industry stakeholders.

Until now, industry research and policy interventions have focused primarily on the statistical underrepresentation of women in key creative roles in film productions (typically writers, directors and producers). However, this focus presumes that the primary instrument for change lies in changing the behaviour of the small number of women in these roles, which will result in amplifying the participation of all women throughout the entire industry. These initiatives typically stipulate that women need support through mentoring or additional “development,” like training programs especially designed for women, and are based on a belief that the underrepresentation of women in the industry results from women’s behaviours rather than men’s (Verhoeven, 2024). Several mentoring programs for the creative industries (Deutscher Kulturrat, 2020) and film industry (Women in Film & TV UK, 2025) have been developed in the last five years, addressing women’s assumed inability to network and to connect with industry executives. To date, this argument has not been empirically examined, nor have the sponsors of these programs conducted formal evaluations. Nevertheless, the effectiveness of this intervention appears to be widely assumed, as it was frequently cited in qualitative interviews conducted for the German Federal Film Board (FFA) (Hochfeld et al., 2017).

Given the film industry’s persistent resistance to various change strategies, scholars and activists have called for a revised approach to the types of data collected and the methods of analysis undertaken – first, to gain a better understanding of the industry’s dynamic structures and, second, to propose different mechanisms to effect change (Verhoeven et al., 2020). This paper adopts the first motivation: To better understand the underlying mechanisms. Specifically, it seeks to examine a central argument concerning gendered industry structures. Our primary research question is therefore:

Do gendered film networks exist?

To address this, we further ask:

How are industry networks constituted and sustained?

Can we identify patterns in how individuals come together within these networks?

Does the educational path play a role in forming these connection patterns?

To answer these questions, we examine the German film industry as a collaboration network in which key creative people (writers, directors and producers) work on projects together. We draw on the approach of Verhoeven et al. (2020), who use Social Network Analysis to understand and analyse gender inequity in order to identify more effective mechanisms for structural change. In this article, we provide an overview of the German film industry, introduce the use of SNA for examining equity within the sector, and outline our methodological approach. We then share findings from our expanded analysis of German film industry data. These findings are contextualized through an interpretation of collaboration net-

works within the German film industry, supplemented by career biographical information on the individual network members.

2. The film industry in Germany: Inequality by the numbers

Each year, approximately 120–140 fiction feature films and an additional 100–120 feature-length documentaries are produced for theatrical release in German cinemas (Filmförderungsanstalt [FFA], 2021, 2025). The market share of tickets sold for German film productions fluctuates between 20 and 30 percent, depending on the specific titles released (Berauer, 2017). As the German film market is primarily confined to German-speaking countries, international distribution for German films is rarely lucrative. Consequently, commercial revenues are limited, and the production of a typical German feature film intended for theatrical release is typically financed primarily through public film funding (Kanzler, 2021). In addition to public subsidies, financing is also secured, to a lesser extent, by television co-production money, with an even smaller portion from a distributor's guarantee. Overall, the majority of German film financing relies on public sources. Annually, around €600 million is made available in Germany through federal, regional, and local film funding agencies (Worschech, 2024). These institutions support a range of objectives – from commercially oriented productions to artistically ambitious works and projects that contribute to regional economic development or tourism. The budget of German theatrical feature films varies: About half of the films have a budget under €2 million, around 15 to 17 percent have a budget over €5 million, and the rest lies between €2 million and €5 million (Prommer & Siegemund, 2022, p. 18). Approximately 90 percent of German feature films released in cinemas receive some form of film funding (in the form of subsidies, loans or tax refunds), with only a few low-budget or film school productions being realized without such support (Kanzler, 2021).

The gender distribution of key creatives – writers, directors, producers – within the German film industry has been the subject of several studies, conducted either from academic perspectives (Prommer & Loist, 2020, 2026) or by film funding bodies (Hochfeld et al., 2017) and industry organizations (Langer, 2019; Prommer & Siegemund, 2022). The most comprehensive of these is *Gender and Film* (Hochfeld et al., 2017), commissioned by the German Federal Film Board (FFA). Unlike earlier studies that focused solely on key creative roles, this study examined gender representation across all creative departments in film production, including directing, screenwriting, producing, cinematography, sound, editing, production design and costume. Its data analysis examined the gender of the heads of department across all German feature and documentary films that screened in cinemas between 2011 and 2015. Findings revealed a significant overrepresentation of men as heads of department in nearly all categories, regardless of film type (Hochfeld et al., 2017). In addition, qualitative interviews conducted for the study highlighted systemic barriers identified by female directors. These included the dominance of closed male networks, persistent gender stereotypes and biases, and opaque decision-making processes – all of which were described as gatekeeping mechanisms within the industry (Hochfeld et al., 2017, p. 56).

The most recent data from 2023 indicates a slight improvement for the year 2019, with women accounting for 31 percent of all directors (up from 29%). However, the sustainability of this progress is questionable as film funding decisions in 2023 continue to reflect significant disparities: Proportionally less funding is allocated to films directed by women (Maier et al., 2024). The overrepresentation of male directors in Germany is comparable to that in other European countries. According to the *Where Are the Women Directors?* report by the European Women's Audiovisual Network, men accounted for 75 to 80 percent of directors in the seven European countries surveyed (Aylett et al., 2016). More recent European data from 2023 show no progress, with women directing only around 20 percent in most European countries (Simone, 2024).

In Germany, only a limited number of policy interventions aimed at improving gender equality have been initiated. These include the introduction of gender-balanced juries at the German Federal Film Funding Board (FFA), the recognition of child care costs as eligible production expenses within funding applications (FFA, 2022), and the development of more transparent reporting on the gender distribution of funding decisions involving the FFA and some of the regional film funders (Medienboard Berlin Brandenburg, 2021). Additional measures include the establishment of mentoring programs for women (Steinbauer, 2019). To date, only one regional funding agency – the MOIN in northern Germany – requires that applicants complete a diversity checklist as part of the funding process (MOIN Filmförderung Hamburg Schleswig-Holstein, 2020). Unlike Austria, which has introduced a financial gender incentive (Österreichisches Filminstitut, 2019), or the United Kingdom with its BFI Diversity standards (BFI, 2026; New Inclusion, 2022), Germany lacks binding gender incentives, quotas or other mandatory control mechanisms. Nevertheless, the topic of inequality in hiring talent remains a constant point of discussion. No film festival, industry conference, or national industry event takes place without raising the issue of gender imbalance in the sector.

3. The network approach to the film industry

3.1. Social Network Analysis (SNA) in the film industry

Films are not made by a single person: they are made by creative teams. Building a creative team in the film industry primarily relies on personal connections in the form of pre-existing professional relationships (Faulkner & Anderson, 1987; C. Jones, 1996; Wreyford, 2015). Because most professionals in the film sector work as freelancers, creative teams are typically assembled on a project-by-project basis. This freelance model, common to many creative industries, results in a high degree of dependence on informal, reputation-based structures to gain employment (Apitzsch, 2017; Blair, 2001; Lutter, 2015). Film workers not only need to maintain strong professional networks to access job opportunities; they also rely heavily on word-of-mouth recommendations, visibility through awards or nominations, and screenings at festivals (Blair, 2001). Crucially, they must build connections with producers and production companies who hire them – who in turn

depend on relationships with commissioning editors and funding agencies to greenlight the projects (Windeler, 2004).

In Germany, only a few large production companies have dedicated departments for evaluating submitted scripts and recruiting talent; therefore, this remains an exception. Most creative teams in the German screen industries are assembled through informal networks. Indeed, Sydow et al. (2002) explicitly describe the German television production sector as a network of organizations and freelance workers (Sydow & Windeler, 2006; Sydow et al., 2002; Windeler, 2004; Windeler et al., 2001). Drawing on qualitative interviews, their research illustrates how these freelance production networks enhance both the flexibility and the operational efficiency of the creative industry.

Adopting a Social Network Analysis (SNA) approach to film industry recruitment can enable researchers to uncover the hidden dimensions of informal collaboration networks by examining the relational structures among individuals working together in production teams. By visualizing and quantifying relationships, SNA can identify key players and roles, reveal collaboration clusters, and detect structural biases in employment patterns. For example, SNA that incorporates an equity lens has shown how small, seemingly inconsequential initial gender inequalities can, over time, wedge people into social network positions and become a basis for cumulative disparities that are increasingly difficult to undo (McPherson & Smith-Lovin, 1993). Organizational Network Analysis (ONA), which focusses specifically on the relationships formed between group members within professional settings (from work teams to board membership), can show how networking, social identities and social capital affect connectivity and workplace success. In their scoping review of this literature, Decker-Tonnesen et al. (2025) identify a dearth of studies relating specifically to men's networks but a rich vein of research assessing women's workplace networks as more peripheral and less connected. Studies on creative industry employment patterns repeatedly show that women are subject to exclusionary network dynamics in these work settings (Grimme et al., 2026). An emerging body of research has begun applying SNA methods to the study of the film industry (Dadlani et al., 2024), with some studies showing that these exclusionary network practices are also important for understanding the role of gender in the formation of film production teams (Jones et al., 2024; Lutter, 2015).

In their comprehensive overview of the existing literature, Dadlani et al. (2024) identify three main trends in SNA film analysis: Research focussed on networks of individuals, research focussed on the connections between films, and research focussed on other links, such as those between commercial companies, or organisations, such as festivals, awards ceremonies, production locations, and so on. Existing work using SNA to analyse the German film industry falls into the latter category. Krätke (2002), for example, examines the location of German filmmaking in order to identify regional production clusters. Similarly, Hoyler and Watson (2018) include the German film industry in a global comparative analysis of film production networks, using SNA to trace spatial and organizational clusters of activity.

Overall, as Dadlani et al. (2024) note, only a limited number of studies have used SNA to examine social inequalities in the film industry from a relational perspective. Among the notable exceptions, Verhoeven et al. (2020) build on the work of Lutter (2015), applying an SNA approach to the Australian, Swedish and German film industries to reveal patterns describing which specific male producers did not work with women. Additionally, they created “what if scenarios” to evaluate the effectiveness of a variety of interventions (“network control mechanisms”) intended to produce more gender equity in these industries. Their research shows that the most effective way to create more equitable conditions for women in the film industry is not necessarily to remove influential men (“key players”), nor to simply add more women, but to ensure that women in the industry are linked to influential men.

Since our main focus lies on the dynamics of team formation, we are specifically interested in the relations between network members in the film industry, and in particular in the role of the film producer, who is largely responsible for recruitment. We therefore adopt a two-step approach: First, we use social network visualizations to identify the most influential producers in the network; second, we add qualitative career-biographical data on the professional career of team members to extend our understanding of patterns underlying the constitution of creative teams. In this way, we combine SNA with qualitative career-biographical data as a form of exploratory data analysis (EDA) (Du Toit et al., 1986). For the biographical career information of industry members, we researched and coded the following variables: Year of birth, place of birth, attendance at film school, other education, specific name of film school, city of the company registration, and city of residence. Since existing data show that women and men attend film schools in equal proportions, we are specifically interested in examining what happens to men and women, respectively, after film school.

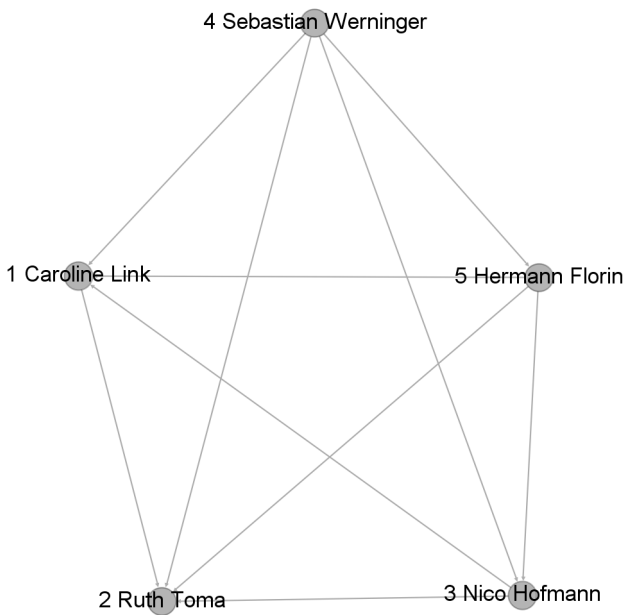
3.2. Network analysis methodology for the film industry

The principle of SNA is based on describing the mathematical graph of social relations in a network in which people are represented by nodes, and if people have a connection, then these nodes are visibly linked by ties, also known as edges. The relationships between nodes can be described using a range of measures that evaluate their levels of connectivity, distribution, and fragmentation. The social networks formed by the relationships between nodes can be modelled and visualized using tools such as Gephi (Bastian et al., 2009). Importantly, Gephi can perform network calculations and offers different layouts to explore the data, such as Force Atlas or Open Ord (Ng & Soo, 2018). Gephi also enables different attributes (colour, size) to different node and edge properties. There are also various analytic methods available to visualize networks. For example, if a *force-directed algorithm* was used to visualize the network, then nodes without edges would repel each other while more closely linked nodes would attract each other (Ng & Soo, 2018).

Applying these graph-based and graphic approaches to the film industry focuses attention on the relational behaviours of people who collaborate on projects.

In the Social Network Analysis of the German film industry that follows, key creative team members, the director, writer and producer, are nodes. They are connected when they work together on film productions. For example, Academy Award-winning director Caroline Link used the screenplay of author Ruth Toma for the film *Der Junge muss an die frische Luft* (*All About Me*). The film was produced by Hermann Florin, Nico Hofmann and Sebastian Werninger. A visualization of these relationships as a social network can be seen in Figure 1.

Figure 1. The social network of the film *Der Junge muss an die frische Luft* (*All About Me*)



Note. All visualizations generated by using Gephi OpenOrd, no overlap algorithm (Prommer & Görland, 2022, p. 111), redesigned for better readability.

In our given example, Caroline Link has collaborated with multiple producers throughout her career, meaning her network includes numerous connections to other individuals. When representing the film industry over an extended period, it is essential to select a layout algorithm that enhances readability and reveals underlying patterns. For example, network visualizations can be adjusted to reflect the strength of relationships – such as by thickening the edges between individuals who have worked together repeatedly over time – or to highlight central figures by displaying nodes with many connections in a larger size. To effectively visualize the complex and overlapping professional relationships that develop over many years in the film industry, we employed the OpenOrd algorithm, previously used by Verhoeven and Palmer (2016), which is well suited for large-scale network data (Martin et al., 2011; Verhoeven, 2017). In the following section, we

analyse the overall network structure of the German production landscape and then zoom in to analyse its structural components in greater detail.

We will focus on the production roles of the producer, screenwriter, and director – commonly referred to as the “key creative team.” This definition is widely used to measure gender inequality in the screen industries. While we acknowledge the significance of other creative roles in film production – such as the director of photography (DOP), editor, sound and costume designers – our analysis adheres to the industry definition of “key creative team” for purposes of constituency and comparability.

It is important to note that our analysis does not approach production networks from a purely statistical perspective. Rather, we use SNA as a visualization tool for exploratory data analysis (EDA), with the aim of uncovering patterns, spotting anomalies, and testing hypotheses (Du Toit et al., 1986). While the structural metrics of the German film networks have already been modelled in previous work (Verhoeven et al., 2020), our approach leverages SNA primarily for visualization, allowing us to examine specific patterns of collaborations. These visual insights are further enriched through integration of qualitative career-biographical data.

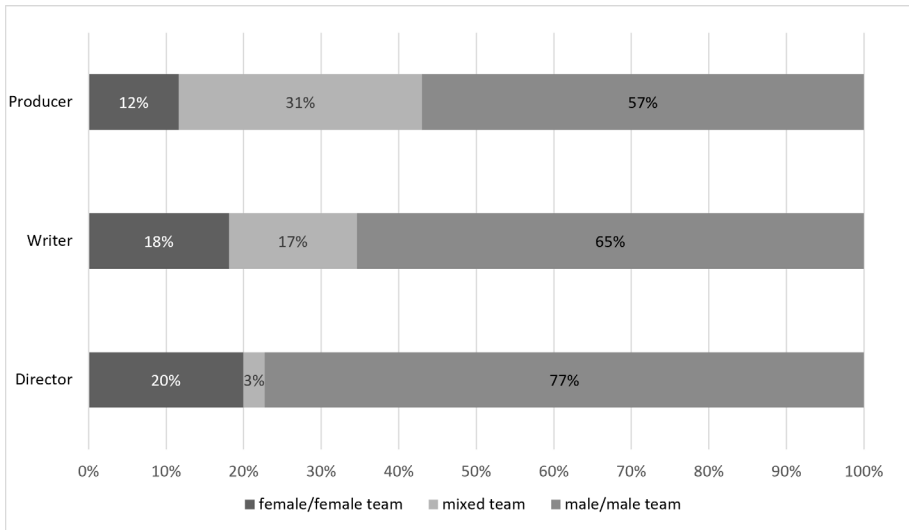
Looking at our given example, we are interested in exploring how the producers Nico Hofmann, Sebastian Werning and Hermann Florin connected to Caroline Link. To answer this question meaningfully across the German film industry, we researched the careers of the most influential producers and the directors and writers they have worked with beginning at their point of entry to the industry.

4. The German film production networks

4.1 Description of data

Our dataset comprises all feature films produced in Germany that premiered in cinemas between 2009 and 2016. This includes all fiction feature films with either majority- or minority-German financing intended for theatrical release, in accordance with the industry definition of a German film. This full sample was compiled from the annual reports of the German Federal Film Board (FFA, 2016). For each film we manually coded the gender of the director, producer, and writer. In addition, we collected the data for distribution, funding, audience numbers, as well as festival screenings and award recognition. This comprehensive dataset provides a detailed overview of the German film industry during the period under study, encompassing a total of 1,139 films.

Figure 2. Gender composition of creative teams in all theatrical feature films released in Germany (2009–2016, $N = 1,139$)



Our full sample of 1,139 German feature films produced between 2009 and 2016 reveals significant gender imbalances in the compositions of key creative teams (producers, directors, and writers). Over half of these films (57%) were produced exclusively by male producers or all-male producer teams. Women producers were involved in 43 percent of the productions, including mixed-gender teams. However, producer teams composed entirely of women account for only 12 percent of all productions. This means that men were involved as producers in approximately nine out of ten feature films, while more than half of the productions had no women producers involved at all.

Similar disparities are evident across other key creative roles: In two-thirds of the films (65%), no women were involved in writing scripts, while men (co-)wrote 82 percent of the scripts. Men helmed 77 percent of these films, while women directed 20 percent, and three percent of the films were directed by mixed-gender teams.

Loist and Prommer (2019) have emphasized the crucial role producers play in creating team diversity and gender equality in production roles. This is clearly reflected in our dataset: When women served as producers, they hired women directors or other female team members more than twice as often (33%) as all-male producer teams did (14%). The same pattern applies to screenwriting, with women writers engaged more than twice as often by women producers (27%) compared to by men producers (12%).

Our data further demonstrates that German women producers face unequal opportunities across multiple dimensions. Their status is not on par with that of their male counterparts. Male producers receive significantly more funding, benefit from better distribution opportunities, and consequently achieve higher au-

dience reach. On average, films produced by men are awarded nearly twice as much funding for their projects (€770,000) as those produced by women (€400,000). Despite one in ten films being produced by all-women production teams, these projects received only seven percent of the total national film funding.

Distribution opportunities are similarly unequal. Films produced by women are released, on average, on 75 screens, whereas films produced by men launch on 118 screens, benefitting from larger marketing budgets and broader audience exposure.

Given that men tend to hire other men, we further analysed gendered participation within creative teams. We found that 450 films (39%) did not include a single woman in any key creative role. Conversely, men were absent from the creative teams in only in 12.4 percent of films. In other words, women are excluded from creative teams three times more often than men.

These findings, which highlight the stark underrepresentation of women in key creative roles, raise pressing questions: What patterns of behaviour govern the formation and structure of creative teams? And do specific professional networks structurally and systemically exclude women from participating in film projects?

4.2 Methodology – Visualizing the production networks

To generate a collaboration network for analysis, the production data was reformatted to track how creative teams were assembled across eight years of German film productions (see Table 1). We compiled a list of individuals and their corresponding gender for all key creative personnel involved in the 1,139 films in our dataset. These film projects employed a total of 2,848 unique individuals, occupying 5,379 key creative roles. Within our sample, women held approximately one-quarter (23%) of the creative positions ($N = 1,231$), while men occupied three-quarters (77%; $N = 4,145$), and 3 individuals identified outside binary gender categories. To enable a detailed analysis of gender representation in the film sector, we manually assigned gender categories “woman,” “man,” or “in another way” – to each unique director, producer, and writer, based on first names and available pronoun information. The average size of a key creative team in German film production is 3.7 members. It is not uncommon for individuals to occupy multiple roles within a single project, such as serving as both screenwriter and director.

Table 1. Overview of feature film production data in Germany (2009–2016).

Fiction feature films ($N = 1,139$)	Germany 2009–2016
Positions	5,379
Unique individuals	2,848
Function director, proportion male	79%
Function screenwriter, proportion male	75%

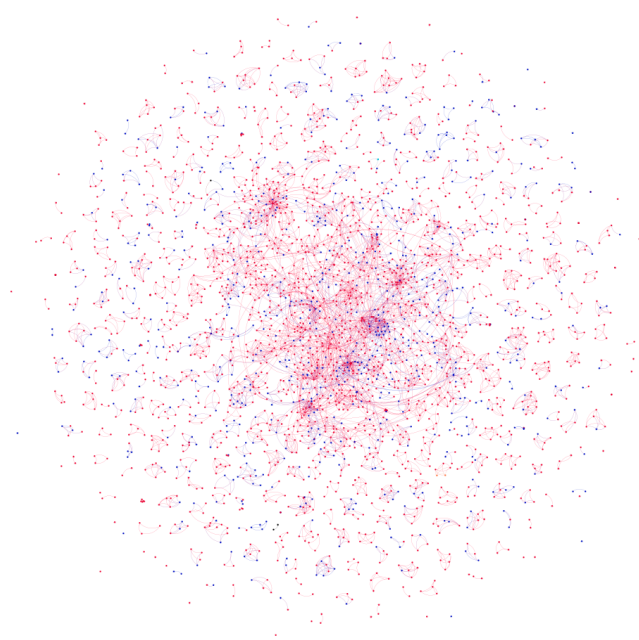
Function producer, proportion male	77%
Producers (unique individuals)	1,446
Average size of creative team	3.7
Share of male producers who worked with 0 women in the years 2009–2016	45%
Share of male producers who worked with 1 woman in the years 2009–2016	30%
Share of films with women in minority role	79%

Our analysis reveals that 45 percent of all male producers in Germany did not work with a single woman in their creative team over the eight-year period covered by our dataset, while an additional 30 percent collaborated with a woman only once. Given that the average size of a creative team is 3.7 members, teams featuring just one woman fall far short of gender parity. In total, this means that 75 percent of male producers in Germany consistently worked in male-dominated teams across an eight-year period. These findings confirm and align with results from previous research applying similar analytical approaches across various national industries, including Germany (Verhoeven et al., 2020).

In our network visualizations of the German film industry, we position the producer as a “source” node, reflecting their formal authority in assembling the creative team – an approach informed by Sydow and Windeler’s conception of producers as central coordinators within project networks (Sydow & Windeler, 2006). To emphasize this dynamic, we constructed a directed network that maps key creative collaborations with reference to the producer’s role.

The network visualisations, created using Gephi, represent individuals as nodes and shared film projects as edges linking them. The following visualizations are designed to illustrate patterns of gendered collaboration. To visualize gender as a variable, we depicted men as red nodes and women as blue nodes. The connecting lines (edges) represent instances of their working collaborations on the same project. For reasons of confidentiality, the names of individual creative team members have been anonymized.

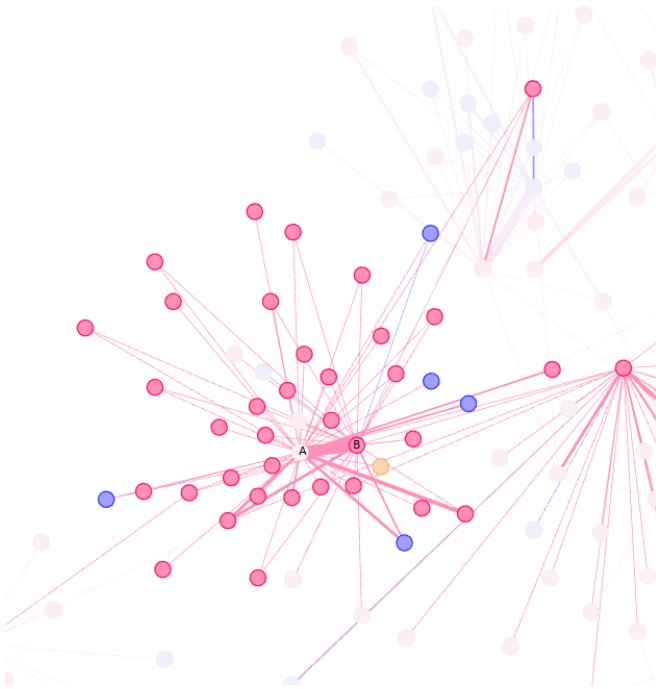
Figure 3. Complete collaboration network: German producers, directors, and writers (2009–2016)



Notes. In the network 2,848 unique individuals are connected via 1,139 films. Red nodes represent men, blue nodes represent women. Edge direction is clockwise from the source node (producer) to the target node (other creatives).

The network visualizations (Fig. 3) clearly demonstrate the dominance of red nodes, representing men, confirming the gender imbalance observed in our quantitative analysis. A closer inspection of the network reveals clusters of male collaborators who have either never worked with a woman or have done so very rarely – visually reinforcing the patterns described in the statistical analysis above. At the periphery of the network, we observe projects involving individuals who participated in only a single film. Isolated nodes typically represent cases where one person held all three key creative roles – producer, writer, and director – on a single production.

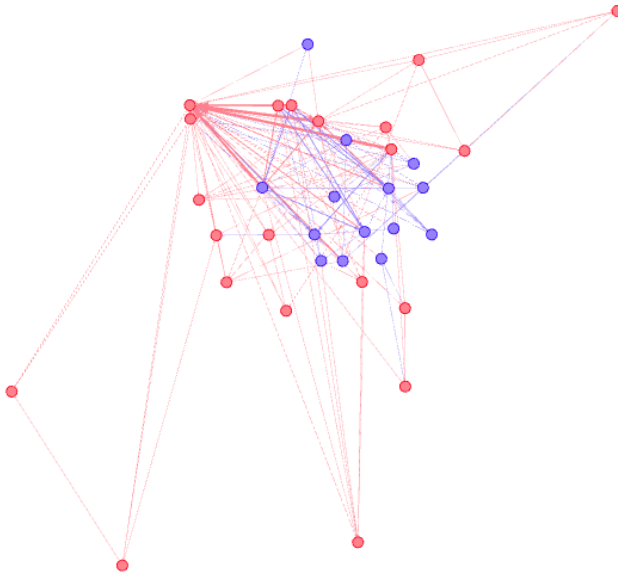
To understand the network in greater depth, we examined a specific segment of the producer collaboration network. Figure 4 highlights a cohesive, male-dominated creative cluster that is representative of broader trends in the data. This cluster is centred around two influential producers (nodes A and B), who co-own the same company. The network includes 37 men and only five women – mostly screenwriters, one co-producer, and no female director. Over the eight-year period covered by our data, this production company did not employ a single female director, and women were hired as screenwriters in just 12 percent of its projects.

Figure 4. Close-up view of a predominantly male producer collaboration network

Notes. Producer collaboration network. In this network, two influential producers (A & B) are connected to 37 men and five women (primarily screenwriters, no directors; women represent 12 percent of the network). The orange node is a nonbinary individual. The line thickness represents the number of joint projects between producers – the thicker the line, the more often producers have worked together on projects.

Another detailed example is shown in Figure 5, illustrating the networks of those very few producers who work with both men *and* women. In this social network, women work as directors, co-producers, and writers. Although two strong producers are also visible, they do not occupy central positions in the network and maintain a greater number of connections to women. This network represents a large film and television production company.

Figure 5. Snapshot of a producer collaboration cluster involving both men and women



Note. The network depicts a team of two producers collaborating with 15 women (42%) and 21 men (58%) within their creative network of a large film and television production company.

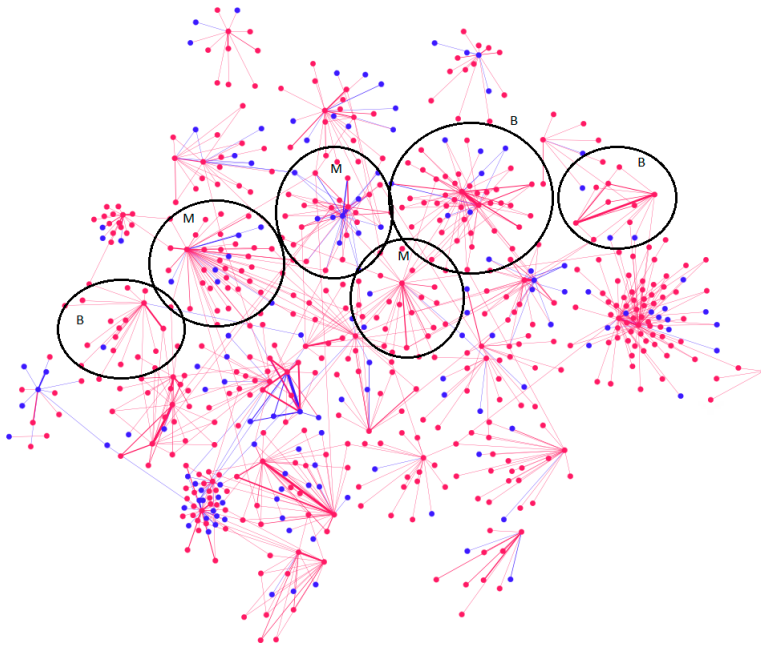
5. The influential producers and their networks

Since our focus is on producers and the formation of their networks, we conducted an even closer and more detailed analysis of 1,466 individual producers. The vast majority (71%, $N = 1,047$) produced only one single film during the study period, while an additional 22 percent ($N = 314$) produced two or three films. We defined the most active producers as those who produced at least one film every two years, meaning four or more films between 2009 and 2016. This group comprises 102 individuals, representing seven percent of all producers in our sample. Among these 102 highly active producers, only 15 individuals produced more than 10 films during the period, and of these, just two were women. The two most prolific producers worked, respectively, on 20 and 23 projects each and both men own their production companies. Our analysis centres on this group of “influential producers,” examining how their predominantly male-dominated networks are structured and maintained.

To explore these networks in greater depth, we collected biographical data about the careers of the 102 most active producers and the creative teams (directors and writers) they worked with across the period of our study. Career biographies were synthesized from company websites, industry publications (such as mediabiz.de), talent agency profiles, and Wikipedia entries. We coded data on educational background (e.g., film school, university, and other pathways), pro-

fessional biography, the city where the production company was registered in 2020, and, when available, the current cities of residence for the directors and writers they worked with. This in-depth individual analysis reveals two major factors shaping network formation among Germany's most influential producers. First, attending the same film school plays a critical role in building professional networks. Second, geographic proximity – captured by the concept of “propinquity” – also emerges as a key driver, with creatives more likely to collaborate when living and working in the same city.

Figure 6. Visualization of the network of the most active influential German producers



Notes. Producer collaboration network of the most active producers ($N = 102$) and their creative teams, based on more than four projects over an eight-year period. Clusters centred in Berlin (B) and Munich (M) are highlighted.

Our analysis identifies three major clusters in Munich, each centred around producers from different generations (teams) who formed professional ties while attending the Munich film school (Hochschule für Fernsehen und Film, HFF München) in the 1980s and late 1990s (Fig. 6). Similarly, we observe three clusters in Berlin: one centred around alumni of the Film University Babelsberg ‘Konrad Wolf’ (formerly Hochschule für Film und Fernsehen Potsdam-Babelsberg, HFF), one around a producing actor, and one other around a filmmaker collective. A striking commonality of all these clusters is the noticeable underrepresentation of women within their networks.

Many of these professional ties, especially in the Munich networks, were established during film school and have been sustained over the course of the graduates' careers, as reflected in the filmographies of the individuals involved. This highlights the crucial role of film schools in affording and shaping enduring industry networks. Prior research confirms that film schools, as pre-industry institutions, often mirror the broader structural inequalities of the film industry (Banks, 2009, 2019). Acknowledging these challenges, the heads of six German film schools signed a joint position paper, "Together for Gender Equality," in 2018. The measures outlined aim to address structural gender discrimination, including promoting equal opportunities, instituting a zero-tolerance policy for gender-based violence and abuse, and adopting a gender-sensitive approach to teaching, particularly in content development and production (Filmuniversität Babelsberg Konrad Wolf et al., 2018).

But why is the film school effect so dominant in the Munich cluster? This can be traced back to German postwar developments (Kaiser & Liecke, 2007). Following World War II and the partition of Berlin, the city lost its status as the centre of the national film industry. The television sector largely relocated to Cologne and Munich, while feature film production became predominantly concentrated in Munich. The establishment of the Munich film school in 1966 further consolidated this development, reinforcing the city's role as a key industry hub. Today, Munich remains home to a dense network of media companies employing around 11,000 people (IHK München, 2026). When correlating years of attendance within the Munich film school clusters, it becomes evident that many individuals began their careers either prior to German reunification in 1989 or in the early 1990s. At that time, entering the film industry in Munich was comparatively easier than in other cities, due to the city's dense concentration of film production and television companies (e.g., Bavaria, Constantin). This context also helps explain the comparatively limited role of the former East German film school, Filmuniversität Babelsberg Konrad Wolf in Potsdam. The profound transformation of industry infrastructures following reunification produced a rupture in the 1990s and gave rise to new institutional and industrial configurations.

While film school affiliations are a major driver in cluster formation, proximity – or propinquity – also plays a significant role. The importance of geographic proximity to key industry hubs such as Berlin and Munich is clearly reflected in the data (Bathelt & Gräf, 2008; Mossig, 2006). Beyond film school-based networks, we identified one cluster of influential producers each in Berlin and Munich that are not anchored in shared educational backgrounds but instead emerge through co-location in major production centres. The centrality of urban hubs in fostering production activity also helps explain the absence of a significant cluster around the Film Academy Baden-Württemberg in Ludwigsburg, a smaller city without a comparable film industry infrastructure. Many of its alumni relocate to Berlin to pursue career opportunities. Similarly, although Cologne constitutes a major hub for television production, it plays a less significant role in feature filmmaking, which is reflected in the relatively small size of the Cologne cluster (Apitzsch & Piotti, 2012). Overall, these findings align with broader patterns of geographic clustering often observed in creative industries (Mossig, 2006).

Although film schools clearly contribute significantly to network formation, they are not the only pathway into the industry. Among the 102 most active producers in our sample, about half attended a film school: A quarter (24) studied at the Munich film school (HFF München), three at Film University Babelsberg in Potsdam and nine at the Filmakademie Baden-Württemberg in Ludwigsburg. The remaining producers entered the industry through alternative routes. Based on our sample, promising pathways into producing include being a son of a successful film family (“nepo-babies”), successful acting careers, experience in distribution companies, or cinema management.

Notably, the more gender-diverse clusters within our network are often associated with producers who attended a variety of film schools or who work across both film and television sectors, or who are affiliated with larger production companies in Germany. In these clusters, women are more frequently represented among the key creative teams. Recruitment strategies at larger, corporate-affiliated companies, such as UFA, indeed appear to support greater gender diversity among key creatives as is proclaimed as a strategic aim by such companies (UFA, 2020).

Anecdotal industry accounts further underscore the role of film school as a foundational moment for building lasting professional networks. A quote from the study *Gender und Film* (Hochfeld et al., 2017, p. 49) captures this dynamic:

I believe that this time at film school is not only important... for learning how to make films, but it's also just as important to build your network. And networking is everything in this industry. So that's just, if nobody knows you, then you're screwed. (Male director, translated from German)

Friendships formed at film school often lead to enduring professional collaborations. However, our network visualizations suggest that this advantage disproportionately benefits men. Women occupy more marginal positions in these networks, with tangible consequences for their careers. Because they are less integrated into these enduring “in-groups”, women cannot rely on them for consistent employment. Instead, they must continually reapply for each new project. By contrast, men benefit from the security of stable, self-reinforcing professional in-groups that offer ongoing employment opportunities. In some cases, this dynamic is so strong that a successful male director, when moving to a larger production company, brings along his former producer as a co-producer, reinforcing these durable male-dominated networks.

6. Limitations

Our study focused on a single national film market; however, as Verhoeven et al. (2020) have shown, there are notable similarities between network structures in the German film industry and those observed in other contexts, such as Australia and Sweden. Whether comparable network patterns exist in film markets with less dominant public funding schemes – such as the United States – remains an open question for future research.

Our analysis is further limited by its exclusive focus on the film industry; television production was not included due to insufficient data, even though many creative professionals work across both sectors. Another limitation concerns the timeframe: The dataset covers an eight-year period, although film projects often spend several years in development before their release, potentially obscuring long-term network dynamics.

For reasons of feasibility, we systematically researched the biographies of only the most active “influential producers” and the directors and writers with whom they collaborated, rather than the full range of production collaborations across the industry. Nevertheless, even within this subset, the pattern of network formation through film school affiliations emerged clearly.

Finally, our gender analysis was basically restricted to binary categories (“man” and “woman”), assigned based on name attribution and pronoun usage, rather than self-identification by individuals within the industry. We only found pronouns for non-binary categories for three individuals in our sample. This methodological constraint reflects data availability rather than conceptual endorsement of binary gender frameworks.

7. Conclusion

This paper has undertaken an exploratory data analysis that combines social network visualisation with detailed career-biography research to generate critical insights into the mechanisms of gender exclusion in film production recruitment. In doing so, it reveals how women’s career trajectories are systemically constrained. Employment in the film industry is inherently precarious, structured around project-based work under unstable conditions and highly sensitive to shifts in a dynamic economic and technological environment. Our findings demonstrate that this precarity is itself gendered: Employment uncertainty is not only experienced differently by men and women, but also structurally organised along these lines.

Our findings raise the pressing question of how such entrenched structures can be changed. A necessary starting point is to rethink existing equity policies that primarily target women’s behaviours. In dynamic industrial settings such as the film industry, regulatory frameworks can play a crucial role in enabling more equitable forms of creative production. However, many existing programs, such as mentoring schemes aimed at improving women’s networking capacities or the creation of women-only networks, fail to address the core issue: The exclusionary dynamics of entrenched, male-dominated networks (Cobb, 2019; Jones et al., 2024).

In this respect, our findings support calls for a shift in focus toward encouraging influential men in the industry to diversify their networks and actively collaborate with women. Some jurisdictions have already taken steps to incentivise such behaviour. For example, both the Austrian Film Institute and the French film funding system have introduced gender incentive schemes (Goodfellow, 2018; Österreichisches Filminstitut, 2019). In Sweden, the Swedish Film Institute implemented gender monitoring and set gender targets, achieving near parity in recent years (Swedish Film Institute, 2018). Although Austria and Sweden are smaller

markets – with roughly one-fourth of Germany’s film production volume – the effectiveness of these initiatives suggests that similar incentive schemes could also be promising for the larger German industry. Given the European film industry’s heavy reliance on public funding, there is both a justification and a responsibility to allocate these funds toward more effective gender equity measures.

Moreover, our research underscores the pivotal role film schools can play in fostering gender-diverse networks from the outset. Both the empirical findings and the career-biographical analysis indicate that professional networks are frequently established during film school and continue to shape career trajectories long after graduation. It is therefore vital that film schools and their educators actively cultivate, and normalize, diverse approaches of creative team building during training, counteract exclusionary patterns of male-dominated network formation, and contribute to the development of more inclusive industry networks (Jones et al., 2026; Kearney & Twenty Female Film Students, 2018).

Finally, this study demonstrates that the often subtle mechanisms of gender exclusion – particularly in the formation of creative teams – can be rendered visible through innovative forms of data analysis and visualization. The integration of aggregated statistical analysis of industry data, relational Social Network Analysis, and detailed career-biographical research provides a more granular understanding of both structural conditions and individual practices, thereby strengthening the basis for meaningful, systemic change.

Generative AI declaration

In the final editing stages of this article’s preparation for publication, one author used ChatGPT for language refinement suggestions for small sections of the text. Outputs were evaluated for accuracy and bias, the authors retained editorial control over meaning and style. All claims and citations derive from sources that were read and interpreted by the authors, who accept full responsibility for its content. No proprietary materials were entered without consent.

References

- Apitzsch, B. (2017). Zwischen Traumberuf und ungewisser Zukunft: Projektarbeitsmärkte in der Film- und Fernsehindustrie. [Between a dream job and an uncertain future: project-based work in the film and television industry.] *KM Magazin*, 119(119), 34–37.
- Apitzsch, B., & Piotti, G. (2012). Institutions and sectoral logics in creative industries: The media cluster in Cologne. *Environment and Planning A: Economy and Space*, 44(4), 921–936. <https://doi.org/10.1068/a44285>
- Aylett, H., Raveney, F., Gore, T., & Prommer, E. (2016). *Where are all the women directors? Report on gender equality for directors in the European film industry*. EWA Women directors in film – Comparative report. European Women’s Audiovisual Network (EWA). https://www.ewawomen.com/wp-content/uploads/2018/09/Complete-report_compressed.pdf
- Banks, M. (2009). Gender below-the-line: Defining feminist production studies. In V. Mayner, M. J. Banks, & J. T. Caldwell (Eds.), *Production studies: Cultural studies of media industries* (pp. 87–98). Routledge.

- Banks, M. (2019). Film schools as pre-industry: Fostering creative collaboration and equity in media production programs. *Media Industries Journal*, 6(1). <https://doi.org/10.3998/mij.15031809.0006.105>
- Bastian, M., Heymann, S., & Jacomy, M. (2009). *Gephi: An open source software for exploring and manipulating networks*. <https://doi.org/10.1609/icwsm.v3i1.13937>
- Bathelt, H., & Gräf, A. (2008). Internal and external dynamics of the Munich film and TV industry cluster, and limitations to future growth. *Environment and Planning A: Economy and Space*, 40(8), 1944–1965. <https://doi.org/10.1068/a39391>
- Berauer, W. (2017). *Filmstatistisches Jahrbuch 2016. [Film Statistics Yearbook 2016]* Nomos.
- BFI. (2026). *BFI diversity standards*. <https://www.bfi.org.uk/inclusion-film-industry/bfi-diversity-standards>
- Blair, H. (2001). ‘You’re only as good as your last job’: The labour process and labour market in the British film industry. *Work, Employment and Society*, 15(1), 149–169. <https://doi.org/10.1177/09500170122118814>
- Cobb, S. (2019). What about the men? Gender inequality data and the rhetoric of inclusion in the US and UK film industries. *Journal of British Cinema and Television*, 17(1), 112–135. <https://doi.org/10.3366/jbctv.2020.0510>
- Conor, B. (2021). *Gender & creativity: Progress on the precipice*. UNESCO. <https://doi.org/10.58337/TTGU8976>
- Dadlani, A., Vo, V., Khemka, A., Harvey, S. T., Kantoro Kyzy, A., Jones, P., & Verhoeven, D. (2024). Leading by the nodes: A survey of film industry network analysis and datasets. *Applied Network Science*, 9(1). <https://doi.org/10.1007/s41109-024-00673-9>
- Decker-Tonnesen, P. L., Chesak, S. S., Walker, L. E., Kohler, K., Phelan, S., Gunnels, M. S., Saliba, K. L., & Bhagra, A. (2025). Role of organizational network analyses to advance workforce inclusion and belonging: A scoping literature review. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1708522>
- Deutscher Kulturrat. (2020, July 15). *Mentoringprogramm für weibliche Führungskräfte im Kulturbereich: Start der 4. Runde. [Mentoring programme for female leaders in the cultural sector: Launch of the fourth round.]* Retrieved from <https://www.kulturrat.de/presse/pressemitteilung/mentoringprogramm-4-runde/>
- Du Toit, S. H. C., Steyn, A. G. W., & Stumpf, R. H. (1986). *Graphical exploratory data analysis*. Springer New York. <https://doi.org/10.1007/978-1-4612-4950-4>
- Faulkner, R. R., & Anderson, A. B. (1987). Short-term projects and emergent careers: Evidence from Hollywood. *American Journal of Sociology*, 92(4), 879–909. <https://doi.org/10.1086/228586>
- FFA. (2016). *Geschäftsbericht 2015 [Annual Report 2015]*.
- FFA. (2021). *Geschäftsbericht 2020 [Annual Report 2020]*.
- FFA. (2022). *Fokus Gender*. FFA Filmförderungsanstalt. <https://www.ffa.de/gender.html>
- FFA. (2025). *Das Kinojahr 2024: Kino- und Filmergebnisse/Filmförderung in Zahlen. Deutscher Film. [The Cinema Year 2024: Cinema and Film Results/Film Funding in Figures. German Film.]*
- Filmuniversität Babelsberg Konrad Wolf, Deutsche Film- und Fernsehakademie Berlin, Filmakademie Baden-Württemberg, Hochschule für Fernsehen und Film München, Internationale Filmschule Köln & Kunsthochschule für Medien Köln (2018). Together for gender equality. <https://rm.coe.int/together-for-gender-equality/168091d93e>
- Flicker, E., & Vogelmann, L. L. (2018). Österreichischer Film Gender Report 2012–2016: *Zentrale Ergebnisse. [Austrian film gender report 2012–2016: Key findings]* <https://filminstitut.at/wp-content/uploads/2023/12/OesterreichischerFilmGenderReport2012-2016.pdf>

- Goodfellow, M. (2018, September 21). France launches gender parity production incentives. *Screen Daily*. Retrieved from <https://www.screendaily.com/news/france-launches-gender-parity-production-incentives/5132855.article>
- Grimme, F., Wittner, B., & Kauffeld, S. (2026). Networking and career success: Analysing social capital in formal female networks. *Journal of Management Development*, 45(2), 226–248. <https://doi.org/10.1108/JMD-01-2025-0017>
- Hochfeld, K., Genz, K., Iffländer, V., & Prommer, E. (2017). *Gender und Film: Rahmenbedingungen und Ursachen der Geschlechterverteilung von Filmschaffenden in Schlüsselpositionen in Deutschland*. [Gender and Film: Context and causes of gender distribution among film professionals in key positions in Germany.] FFA-Filmförderungsanstalt. [https://www.ffa.de/files/ffa/gender/Studie percent20GENDER percent20UND percent20FILM.pdf](https://www.ffa.de/files/ffa/gender/Studie%20GENDER%20UND%20FILM.pdf)
- hooks, b. (2015). *Black looks: Race and representation*. Routledge.
- Hoyler, M., & Watson, A. (2018). Framing city networks through temporary projects: (Trans)national film production beyond ‘Global Hollywood’. *Urban Studies*, 56(5), 943–959. <https://doi.org/10.1177/0042098018790735>
- IHK München (2026, January 9). *Medien*. Retrieved from <https://www.ihk-muenchen.de/politik/interessenvertretung/medien/>
- Jones, C. (1996). Careers in project networks: The case of the film industry. In M. B. Arthur (Ed.), *The boundaryless career: A new employment principle for a new organizational era* (pp. 58–75). Oxford University Press.
- Jones, P., Verhoeven, D., Dadlani, A., & Zemaityte, V. (2024). She must be seeing things! Gender disparity in camera department networks. *Social Networks*, 76, 120–134. <https://doi.org/10.1016/j.socnet.2023.09.004>
- Jones, P., Verhoeven, D., & Espinosa-Rada, A. (2026). The role of film school collaboration in film industry network formation: An empirical analysis of data from the Australian Film, Television and Radio School (AFTRS). *Journal of Cultural Analytics*. Advance online publication. <https://doi.org/10.22148/jca.1173>
- Kaiser, R., & Liecke, M. (2007). The Munich feature film cluster: The degree of global integration and explanations for its relative success. *Industry & Innovation*, 14(4), 385–399. <https://doi.org/10.1080/13662710701524031>
- Kanzler, M. (2021). *Fiction film financing in Europe: A sample analysis of films released in 2019*. <https://rm.coe.int/fiction-film-financing-in-europe-2021-edition/1680a57229>
- Kearney, M., & Twenty Female Film Students. (2018). Melting the celluloid ceiling: Training girl filmmakers, revolutionizing media culture. In M. G. Blue & M. C. Kearney (Eds.) *Mediated girlhoods* (pp. 213–232). Peter Lang.
- Krätke, S. (2002). Network analysis of production clusters: The Potsdam/Babelsberg film industry as an example. *European Planning Studies*, 10(1), 27–54. <https://doi.org/10.1080/09654310120099254>
- Langer, J. (2019). *GENDER.DOK Untersuchung zur Genderverteilung im Bereich Regie bei Reportagen, Dokumentationen und Kinodokumentarfilmen*. [GENDER.DOK Study on gender distribution in directing roles for news reports, documentaries and feature-length documentaries.] <https://media02.culturebase.org/data/docs-ag-dok/GENDER.DOK.pdf>
- Lauzen, M. M. (2021). *The Celluloid ceiling report: Behind-the-scenes employment of women on the top U.S. films of 2020*. https://womenintvfilm.sdsu.edu/wp-content/uploads/2021/01/2020_Celluloid_Ceiling_Report.pdf
- Liddy, S. (2020). *Women in the international film industry: Policy, practice and power*. Palgrave Macmillan.

- Linke, C., & Prommer, E. (2021). From fade-out into spotlight: An audio-visual character analysis (ACIS) on the diversity of media representation and production culture. *Studies in Communication Sciences*, 21(1), 145–161. <https://doi.org/10.24434/j.scoms.2021.01.010>
- Loist, S. (2018). Gendered media industries: Argumente für eine geschlechtergerechte und diverse Filmindustrie. [Gendered media industries: Arguments for a gender-equal and diverse film industry.] *Navigationen*, 18(2), 135–158.
- Loist, S., & Prommer, E. (2019). Gendered production culture in the German film industry. *Media Industries Journal*, 6(1). <https://doi.org/10.3998/mij.15031809.0006.106>
- Loist, S., Verhoeven, D., Eikhof, D. R., Prommer, E., Ehrich, M. E., Jones, P., Guyan, K., Coles, A., Radziwill, S., & Dadlani, A. (2024). *Re-framing the picture: An international comparative assessment of gender equity policies in the film sector*. <https://doi.org/10.60529/390>
- Lutter, M. (2015). Do women suffer from network closure? The moderating effect of social capital on gender inequality in a project-based labor market, 1929 to 2010. *American Sociological Review*, 80(2), 329–358. <https://doi.org/10.1177/0003122414568788>
- Maier, T., Amtsberg, L.-B., & Radziwill, S. (2024). *Achter Regie-Diversitätsbericht des Bundesverbandes Regie e.V. (BVR) für die Jahre 2022 und 2023: Genderanalyse zur Regievergabepraxis in deutschen fiktionalen Primetime-Programmen von ARD, ZDF, RTL, SAT.1, Pro7 und VOX sowie der Streaminganbieter und im deutschen Kinospielefilm*. [Eighth Diversity Report on Directing by the Bundesverband Regie e.V. (BVR) for the years 2022 and 2023: Gender analysis of directing commissioning practices in German primetime fiction programmes on ARD, ZDF, RTL, SAT.1, Pro7 and VOX, as well as among streaming providers and in German feature films] Bundesverband Regie e.V., Berlin.
- Martin, S., Brown, W. M., Klavans, R., & Boyack, K. W. (2011). *OpenOrd: An open-source toolbox for large graph layout*. <https://doi.org/10.1117/12.871402>
- McPherson, M., & Smith-Lovin, L. (1993). You are who you know: A network perspective on gender. In P. England (Ed.), *Theory on gender/feminism on theory* (pp. 223–241). Aldine.
- Medienboard Berlin Brandenburg (2021). *MBB Tätigkeitsbericht 2020*. [MBB Annual Report 2020] https://www.medienboard.de/fileadmin/user_upload/pdf/Publikationen/MBB_Taetigkeitsbericht_2020.pdf
- MOIN Filmförderung Hamburg Schleswig-Holstein. (2020). *Diversity Checklist der MOIN Filmförderung* [Diversity Checklist of the MOIN Film Fund]. https://www.moin-filmfoerderung.de/de/ueber_die_filmfoerderung/diversity-checklist-filmfoerderung.php
- Mossig, I. (2006). *Netzwerke der Kulturökonomie: Lokale Knoten und globale Verflechtungen der Film- und Fernsehindustrie in Deutschland und den USA*. [Networks in the cultural economy: Local hubs and global interconnections in the film and television industry in Germany and the USA.] transcript.
- Mulvey, L. (1975). Visual pleasure and narrative cinema. *Screen*, 16(3), 6–18. <https://doi.org/10.1093/screen/16.3.6>
- New Inclusion. (2022). *Review of the BFI Diversity Standards: A Summary of Industry Perspectives & Recommendations*. BFI. <https://core-cms.bfi.org.uk/media/16482/download>
- Ng, A., & Soo, K. (2018). *Data Science – Was ist das eigentlich?! Algorithmen des maschinellen Lernens verständlich erklärt*. [Data science – What is it, anyway?! Machine learning algorithms explained in simple terms] Springer. <https://doi.org/10.1007/978-3-662-56776-0>

- Österreichisches Filminstitut. (2019). *Gender incentive*. <https://filminstitut.at/foerderung/gender-incentive>
- Prommer, E., & Görland, S. O. (2022). Methodische Herausforderungen von Data-Mining, Autocoding und sozialer Netzwerkanalyse: Das Beispiel des Gender-Gaps in der Filmproduktion in Deutschland. [Methodological challenges of data mining, autocoding and social network analysis: The example of the gender gap in film production in Germany.] In C. Lohmeier & T. Wiedemann (Eds.), *Datenvielfalt in kommunikationswissenschaftlichen Forschungskontexten* (pp. 105–125). Springer VS. https://doi.org/10.1007/978-3-658-36645-2_6
- Prommer, E., & Loist, S. (2020). Where are the female creatives? The status quo of the German screen industry. In S. Liddy (Ed.), *Women in the international film industry: Policy, practice and power* (pp. 43–60). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-39070-9_3
- Prommer, E., & Loist, S. (2026). Film industry: Production culture with gender bias. In J. Dorer, B. Geiger, B. Hipfl, & V. Ratković (Eds.), *Handbook of media and gender: Perspectives and findings of feminist communication and media studies* (pp. 499–511). Palgrave Macmillan. https://doi.org/10.1007/978-3-658-45789-1_38
- Prommer, E., & Siegemund, S. (2022). *Siebter Regie-Diversitätsbericht des Bundesverbandes Regie e.V. (BVR) für die Jahre 2019 und 2020. Genderanalyse zur Regievergabepaxis in deutschen fiktionalen Primetime-Programmen von ARD, ZDF, RTL, SAT.1, Pro7 und VOX sowie der Streaminganbieter und im deutschen Kinospielefilm*. Im Auftrag des Bundesverbandes Regie e.V. [Seventh Diversity Report on Directing by the Bundesverband Regie e.V. (BVR) for the years 2019 and 2020. Gender analysis of directing commissioning practices in German primetime fiction programmes on ARD, ZDF, RTL, SAT.1, Pro7 and VOX, as well as among streaming providers and in German feature films]
- Ross, K., & Padovani, C. (Eds.). (2017). *Gender Equality and the Media: A Challenge for Europe*. Routledge.
- Screen Australia. (2015). *Gender matters: Women in the Australian screen industry*. Screen Australia. <https://www.agec.org.au/wp-content/uploads/2018/09/Gender-Matters-Women-in-the-Australian-Screen-Industry-2017.pdf>
- Simone, P. (2024). *Female professionals in European film production: 2023 edition*. <https://rm.coe.int/female-professionals-in-european-film-production-eao-report/1680ad13c5>
- Smith, S. L., Choueiti, M., Pieper, K., Case, A., Yao, K., & Vranich, E. (2020). *Inequality in 1,300 popular films: Examining portrayals of gender, race/ethnicity, LGBTQ & disability from 2007 to 2019*. USC Annenberg Inclusion Initiative.
- Smith, S. L., Choueiti, M., Yao, K., Clark, H., & Pieper, K. M. (2020). *Inclusion in the director's chair: Analysis of director gender & race/ethnicity across 1,300 top films from 2007 to 2019*. USC Annenberg Inclusion Initiative. <http://assets.uscannenberg.org/docs/aai-inclusion-directors-chair-20200102.pdf>
- Steinbauer, A. (2019, March 4). Weiblicher, jünger, erfolgreicher. [More feminine, younger, more successful.] *Süddeutsche Zeitung*. Retrieved from <https://www.sueddeutsche.de/medien/zdf-weiblicher-juenger-erfolgreicher-1.4353999>
- Swedish Film Institute. (2018). *The money issue: Gender equality report 2018*. <https://www.filminstitutet.se/globalassets/2-fa-kunskap-om-film/analys-och-statistik/publications/other-publications/sfi-gender-equality-report-2018---lowres.pdf>
- Sydow, J., & Windeler, A. (2006). Neue Organisationsformen in der Medienökonomie. [New forms of organisation in the media industry.] In K.-D. Altmeppen & M. Karmašin (Eds.), *Medien und Ökonomie: Band 3: Anwendungsfelder der Medienökonomie*

- (pp. 47–60). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90195-4_3
- Sydow, J., Windeler, A., & Wirth, C. (2002). Markteintritt als kollektiver Netzwerkeintritt: Internationalisierung der Fernsehproduktion in unreife Märkte. [Market entry as collective network entry: The internationalisation of television production in immature markets.] *Betriebswirtschaft*, 62(5), 459–473.
- UFA. (2020). *Diversity*. <https://www.ufa.de/die-ufa/verantwortung/diversity>
- Verhoeven, D. (2017). *The “Gender Offender” analysis: How and why we did it (Part two)* | *Kinomatics*. <https://kinomatics.com/posts/the-gender-offender-analysis-how-and-why-we-did-it-part-two/>
- Verhoeven, D. (2024, August 26). Beyond ‘one and done’: Achieving gender equity in the film industry depends on more than entry programs. *The Conversation*. Retrieved from <https://theconversation.com/beyond-one-and-done-achieving-gender-equity-in-the-film-industry-depends-on-more-than-entry-programs-232553>
- Verhoeven, D., Musial, K., Palmer, S., Taylor, S., Abidi, S., Zemaityte, V., & Simpson, L. (2020). Controlling for openness in the male-dominated collaborative networks of the global film industry. *PLOS ONE*, 15(6). <https://doi.org/10.1371/journal.pone.0234460>
- Verhoeven, D., & Palmer, S. (2016, November 15). Women aren’t the problem in the film industry, men are. *The Conversation*. Retrieved from <http://theconversation.com/women-arent-the-problem-in-the-film-industry-men-are-68740>
- Windeler, A. (2004). Organisation der TV-Produktion in Projektnetzwerken: Zur Bedeutung von Produkt- und Industriespezifika. [Organising TV production in project networks: The importance of product and industry-specific factors.] In J. Sydow & A. Windeler (Eds.), *Organisation der Content-Produktion* (pp. 55–76). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80413-6_4
- Windeler, A., Lutz, A., & Wirth, C. (2001). Netzwerksteuerung durch Selektion: Die Produktion von Fernsehserien in Projektnetzwerken. [Network control through selection: The production of television series in project networks.] *Montage AV*, 10(1), 91–124. <https://doi.org/10.25969/mediarep/94>
- Women in Film & Television UK. (2026). *Mentoring*. <https://www.wftv.org.uk/mentoring>
- Worschech, R. (2024, April 26). Filmförderung: Noch viel zu tun. [Film funding: There is still a lot to be done.] *Epd Film*. Retrieved from <https://www.epd-film.de/meldungen/2024/filmfoerderung-noch-viel-zu-tun>
- Wreyford, N. (2015). Birds of a feather: Informal recruitment practices and gendered outcomes for screenwriting work in the UK film industry. *The Sociological Review*, 63, 84–96. <https://doi.org/10.1111/1467-954X.12242>

Providing Ethical Orientation amid Communicative Conflict



Zwischen Hassrede, Framing und generativer Künstlicher Intelligenz Medien und Sprache aus ethischer Perspektive

Edited by Dr. Ingrid Stapf, Prof. Dr. Jessica
Heesen and Laura Martena, M.A.

2026, 338 pp., pb., € 89.00

ISBN 978-3-7560-4020-9

E-Book 978-3-7489-7032-3

(Kommunikations- und Medienethik, vol. 24)

In German

Public communication is essential for moral negotiation processes and democratic understanding. However, the way language is used has itself become a key source of moral and political conflict. This volume presents the current discourse on language use from a media-ethics perspective. Drawing on empirical stu-

dies and media-ethics theory, it examines language-related conflicts, digital communication practices, and AI-supported information processes in terms of their social consequences, with the aim of providing guidance for a democratic culture of communication.

Available digitally at [inlibra.com](https://www.inlibra.com)

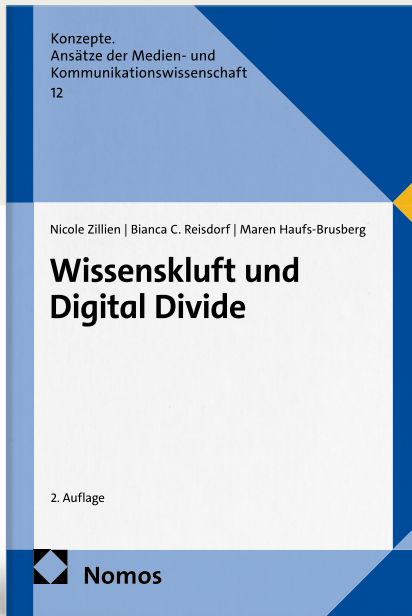
Available in bookstores or free of shipping costs at [nomos-shop.de](https://www.nomos-shop.de)

Returns are at the addressee's risk and expense. All prices include VAT



Nomos

A Closer Look at Media and Inequality



Wissenskluft und Digital Divide

By Prof. Dr. Nicole Zillien, Prof. Dr. Bianca C. Reisdorf, PhD and Maren Haufs-Brusberg, M.A.

2nd, comprehensively updated edition 2026, 98 pp., pb., € 19.90

ISBN 978-3-7560-3053-8

E-Book 978-3-7489-5363-0

(Konzepte. Ansätze der Medien- und Kommunikationswissenschaft, vol. 12)

In German

The second, comprehensively updated edition offers an introduction to the fundamentals, methods and social contexts of knowledge gap and digital divide research. These approaches are based on growing inequalities resulting from the spread of (digital) media, especially in the age of algorithmic decision-making systems and AI. From the thesis of the

growing knowledge gap that emerged in the 1970s to the beginning of digital divide research in the 1990s to the current analysis of digital inequalities, the development history of the approach is outlined based on relevant research results. The volume is aimed at students and lecturers of communication and social sciences.

Available digitally at [inlibra.com](https://www.inlibra.com)

Available in bookstores or free of shipping costs at [nomos-shop.de](https://www.nomos-shop.de)

Returns are at the addressee's risk and expense. All prices include VAT



Nomos