SC|M

Studies in Communication and Media

EDITORIAL

(Re)Establishing quality criteria for content analysis: A critical perspective on the field's core method
Editorial to the Special Issue

Alte und neue Qualitätskriterien für die Inhaltsanalyse: Eine kritische Perspektive auf die zentrale Methode der Kommunikationswissenschaft
Editorial zum Sonderheft

Mario Haim, Valerie Hase, Johanna Schindler, Marko Bachl & Emese Domahidi

Mario Haim (Prof. Dr.), LMU Munich, Department of Media and Communication, Akademiestr. 7, 80799 Munich, Germany. Contact: haim@ifkw.lmu.de
Valerie Hase (Dr.), LMU Munich, Department of Media and Communication, Akademiestr. 7, 80799 Munich, Germany. Contact: valerie.hase@ifkw.lmu.de
Johanna Schindler (Dr.), LMU Munich, Department of Media and Communication, Akademiestr. 7, 80799 Munich, Germany. Contact: johanna.schindler@ifkw.lmu.de
Marko Bachl (Prof. Dr.), Freie Universität Berlin, Institute for Media and Communication Studies, Garystr. 55, 14195 Berlin, Germany. Contact: marko.bachl@fu-berlin.de
Emese Domahidi (Prof. Dr.), Technische Universität Ilmenau, Institute of Media and Communication Science, Ehrenbergstr 29, 98693 Ilmenau. Contact: emese.domahidi@tu-ilmenau.de

EDITORIAL

(Re)Establishing quality criteria for content analysis: A critical perspective on the field's core method

Editorial to the Special Issue

Alte und neue Qualitätskriterien für die Inhaltsanalyse: Eine kritische Perspektive auf die zentrale Methode der Kommunikationswissenschaft

Editorial zum Sonderheft

Mario Haim, Valerie Hase, Johanna Schindler, Marko Bachl & Emese Domahidi

Abstract: Content analysis is one of the core methods of communication science. However, it is currently confronted with several challenges, such as the influx of procedures, data, and measurements emerging from computational methods. To understand how communication science adapts its methods while simultaneously reassuring their ongoing functionality, the six contributions in this Special Issue focus on (re)established quality criteria for content analysis. They showcase the fact that while manual content analysis (and human coders) is still at the core of our methodology, traditional quality criteria are being reinterpreted and approximated, often in light of open science practices and computational text analysis. Therefore, we call for further reflection on conceptual clarity and methodological approaches related to traditional quality criteria (validity, reliability), how they may be reestablished (reproducibility, robustness, and replicability), as well as criteria that have recently come into focus (e.g., ethics). By bringing together leading scholars in this Special Issue, we aim to contribute to moving content analysis forward as a method based on insights from both inside and outside our discipline.

Keywords: Content analysis, quality criteria, validity, reliability, computational methods, interdisciplinarity, methodology.

Zusammenfassung: Die Inhaltsanalyse stellt eine zentrale Methode der Kommunikationswissenschaft dar. Doch sie sieht sich gegenwärtig mit mehreren Herausforderungen konfrontiert, etwa neuen Verfahren, Daten oder Messungen aus dem Bereich der "Computational Methods." Ziel dieses Themenhefts ist es, methodische Fortschritte und Anpassungsprozesse der Inhaltsanalyse zu diskutieren. Die sechs Beiträge konzentrieren sich entsprechend auf (neu) etablierte Qualitätskriterien und zeigen, dass die manuelle Inhaltsanalyse (sowie menschliche Kodierende) nach wie vor den Kern unserer Methodologie bilden. Die Beiträge illustrieren aber auch, dass traditionelle Qualitätskriterien neu interpretiert und angeglichen werden, oftmals als Konsequenz einer zunehmend präsenten Open-Science-Kultur sowie automatisierter Verfahren. Das Themenheft unterstreicht dabei die Bedeutung weitergehender Überlegungen zur Konzeption und Messung traditioneller (Validität, Reliabilität), neu zu etablierender (Re-

produzierbarkeit, Robustheit und Replizierbarkeit) sowie neuer (z. B. Ethik) Qualitätskriterien. Diese und die nächste Ausgabe des Themenheftes bringen hierfür Erkenntnisse innerhalb und außerhalb unserer Disziplin zusammen, um die Inhaltsanalyse als Methode voranzubringen.

Schlagwörter: Inhaltsanalyse, Qualitätskriterien, Validität, Reliabilität, automatisierte Methoden, Interdisziplinarität, Methodenforschung.

1. Introduction

Content analysis is one of communication science's central methods. While the discipline certainly includes a great variety of methods, content analysis is not only one of the main methods that communication scholars regularly employ (Walter et al., 2018) but also the only method originally developed by our discipline (Loosen & Scholl, 2012). As such, content analysis is defined by a long tradition of established procedures, data sources, measurements, and quality criteria.

However, content analysis is currently confronted with several challenges. For one, more and more diverse data from various (online) channels have become relevant, which has introduced challenges for sampling and analyzing relevant data (Jünger et al., 2022; Mahl et al., 2023; Schatto-Eckrodt, 2022). The same applies to new methods for collecting and analyzing data for content analysis, often via automated means (for overviews, see Haim, 2023; Jünger & Gärtner, 2023; van Atteveldt et al., 2022). As such, procedures, data, measurements, and quality criteria from other disciplines, particularly computer science, are increasingly being incorporated into communication science (Bachl & Scharkow, 2017; Baden et al., 2022; Günther & Quandt, 2016; Hase et al., 2023). Increasing interlinkages with these more technical disciplines have also fostered debates about epistemological shifts in the field (Geise & Waldherr, 2021; Helles & Ørmen, 2020; more generally, see Kitchin, 2014).

Given the influx of new procedures, data, measurements, and quality criteria, communication science needs to stay flexible by adapting its methods while simultaneously ensuring their ongoing functionality. To diagnose how well communication science is doing both, we focus on a single yet integral element of methods: quality criteria as criteria used to ensure the quality of scientific research and to differentiate it from nonscientific approaches.

The idea for this Special Issue arose from the 2022 Annual Conference of the German Communication Association's Method Division at LMU Munich. The goal of the conference and the related Special Issue was to move forward and nurture content analysis as a central method of the discipline in times of change.

2. State of the art

Currently, content analysis faces several methodological debates concerning existing and emerging quality criteria (Casas & Webb Williams, 2022; Krippendorff, 2018, 2021; Lacy et al., 2015; Sommer et al., 2014), similar to other social science methods (RatSWD, 2023). These criteria seem to lack conceptual clarity (Freiling et al., 2021) and consensus, especially in light of computational methods changing existing or introducing potentially new criteria (Domahidi et al., 2019; Geise & Waldherr, 2021; Haim, 2021). Moreover, and often related to new methods, scholars criticize the fact that, to date, "everyone brings the practices and standards from their original field" (Theocharis & Jungherr, 2021, p. 12) to the table. As such, "a lack of currently established standards […] can jeopardize the scholarly scrutiny which is essential in assuring additive science and replicability" (van Atteveldt et al., 2019, p. 3). This is certainly true for computational methods, which are the subject of debate in many scholarly and general methodological discussions (Theocharis & Jungherr, 2021; van Atteveldt et al., 2019). These discussions may also extend to more traditional approaches, such as qualitative or quantitative manual content analysis (Krippendorff, 2021; Lacy et al., 2015).

Across disciplines and methods, different quality criteria coexist (for related debates, see LeBel et al., 2018; RatSWD, 2023). In our Call for Paper for this Special Issue, we initially focused on five criteria (validity, reliability, reproducibility, replicability, and robustness) that, from our guest editors' perspective, seemingly required fresh perspectives to diagnose the state of content analysis and, as such, the field.

A seminal quality criterion, *validity*, is crucial to content analysis, as it describes, among other aspects, whether the results correspond to some empirical truth (e.g., Krippendorff, 2018; Song et al., 2020). Common to communication science is the understanding that this empirical truth is what humans make of the content to be analyzed. As a quality criterion, validity then defines how to approximate this correspondence. Agreement on the criterion of validity has been debated in relation to qualitative approaches (Dutta et al., 2020) and challenged, in particular, by the rise of automated approaches. Can we use automated content analysis to measure latent theoretical concepts (Baden et al., 2022), how can such methods be validated (Grimmer & Stewart, 2013), and can manually annotated texts serve as sufficient "ground truth" (Song et al., 2020)?

Validity is closely linked to *reliability*, which allows researchers to quantify whether repeated measures of the same data yield similar results. In other words, if multiple human coders link the given content to the same empirical truth, then it is presumed to be both reliable (they all did it the same way) and valid (they all reached the same interpretation of the content, so that must be *the* truth). Here, recent debates are concerned with how to combine coding from different coders and how to correct for errors emerging from disagreements, including sufficient thresholds for reliability (Bachl & Scharkow, 2017; Geiß, 2021; Krippendorff, 2021; TeBlunthuis et al., 2023).

While validity and reliability are well-established quality criteria (Krippendorff, 2018; Potter & Levine-Donnerstein, 1999), recent years have opened up interpretations of reliability that go beyond the traditional understanding of inter-coder agreement. These include *reproducibility* (i.e., if others reach the same results based on the same data and methods), *replicability* (i.e., whether the results hold when applying the same methods to different data), and *robustness* (i.e., whether the results change when using different methods for the same data). In particular, the former two have gained in importance in line with open science principles (Dienlin et al., 2021; Freiling et al., 2021). While reproducibility, replicability, and robustness may constitute new ways of thinking about reliability as a quality criterion, they lead to new debates. Not only may we have to reconsider what these criteria mean conceptually in light of automated methods (Schoch et al., 2023), but we may also lack research on how to apply them in practice. Content analyses, for example, are less often replicated than studies using other methods (Keating & Totzkay, 2019). Similarly, research from computational linguistics (Wieling et al., 2018) and sociology (Nelson, 2019) indicates that automated content analysis may not necessarily be easier to reproduce, despite automation through scripts seemingly promising such. Related to robustness, researchers' decisions on data cleaning, processing (Pipal et al., 2023; Wilkerson & Casas, 2017), and analysis methods (Nelson et al., 2021) may impact the results of content analysis, something that is often discounted.

Given current debates about validity, reliability, reproducibility, replicability, and robustness as (re)established criteria, our Special Issue is concerned with the following question: How have established quality criteria related to content analysis changed – and, given increasing interlinkages with other disciplines in communication science, have new quality criteria been established?

3. (Re)Establishing quality criteria

In the first contribution to this Special Issue, Oschatz, Sältzer, and Stier argue that while the conventions of appropriate thresholds for reliability as a prerequisite for validity have gained discursive attention with the increasing popularity of automated content analysis, social scientists will not be able to abandon manually coded data. The authors claim that (semi-)supervised machine learning depends on high-quality – reliable and valid – human coding as the most important factor for adequate training and subsequent prediction of machine learning algorithms that aim to classify the (empirical) meaning of text. Oschatz and colleagues consider two sequential sources of threats to reliability and, thus, validity that may affect the quality of training data: (non-)systematic errors in human annotations and curation strategies for dealing with disagreement. They then set out to simulate how these error sources may affect validity to establish standards for automated content analysis.

Similarly addressing tensions between reliability and validity, the second contribution to this Special Issue by Baden, Boxman-Shabtai, Tenenboim-Weinblatt, Overbeck, and Aharoni challenges the common assumption that there is a definite and unique way to classify any instance of a measured variable correctly. The authors argue that there are content ambiguities that can cause meaning multiplicity. In such cases, coding disagreement should not be understood as measurement error but reflects the true properties of the coded content. The authors therefore introduce the notion of "valid disagreement," a form of reliable disagreement about the correspondence between content and empirical truth that must be distinguished from threats to validity.

In the third contribution to this Special Issue, Niemann-Lenz, Dittrich, and Scheper offer another take on reliability and validity by examining the quality of coding when comparing student coders and crowdworkers. They present empirical data on how coder characteristics influence coding quality, both for manifest and latent categories. Their comparison shows sufficient validity for more manifest categories but not for more latent categories, as well as a slightly higher reliability of student coders compared to crowdworkers.

In the fourth contribution to this Special Issue, Wiedicke discusses how ethics, as a quality criterion that has undergone much discussion recently but is rarely reflected upon in the context of content analysis, is closely linked to validity and overall research quality. Research ethics have gained greater prominence in recent years, as contemporary content analysis faces challenges that range from the use of automated approaches to data analysis, potential biases in platform data, and insufficient informed consent in online data collection. Drawing on the principles-based approach (i.e., respect for a person's autonomy, beneficence, nonmaleficence, and justice), Wiedicke presents a systematic overview of and possible pathways for the ethical challenges and potential dilemmas related to methods and ethics arising from contemporary content analysis.

In the fifth contribution to this Special Issue, Rieger, Yachenko, Ruckdeschel, von Nordheim, Kleinen-von-Königslöw, and Wiedemann showcase the implementation and evaluation of pre-trained language models in content analysis. Their contribution, which will be published in the next issue of SCM due to space restrictions, offers an overview of the potential and impediments of such models. The authors address these challenges by using a multilingual transformer model with adapters and few-shot learning to showcase a parameter-efficient, easily shareable, and well-performing approach. Beyond its specific case and candidate models, the article is a best-practice example for evaluating and reporting an innovative method's validity, reliability, replicability, reproducibility, and even robustness.

Finally, in the sixth contribution to this Special Issue, Chan, Freudenthaler, and Müller address the issue of validity for framing, a particular motif of content analysis. The coding of frames, not least by computational means, has a long tradition of being criticized for inaccuracy, lack of sensible conceptualization, and difficulties in implementation, particularly vis-à-vis the coding of topics. The authors present a synthetic dataset for evaluating frames identified in multi-topical news content to benchmark manual coding and various automated and semi-supervised methods. Somewhat ironically, however, Chan and colleagues provide evidence that generic frame identification using both manual coding and automated methods might not be accurate (enough). Again, this contribution will be part of SCM's next issue.

4. Moving on

From these contributions, as well as from our reflections on the mismatch between a priori specified and a posteriori submitted quality criteria, we conclude with four main diagnoses on the status of content analysis as a central method in communication science and how to move forward. We consider these to be both a positive testimony to the current state of content analysis and crucial when moving on and continuing to nurture the field's key method.

First, content analysis is alive and in good health. The number of submissions and lively discussions both at the conference and in the reviews of these submissions clearly indicate the strong ambition to investigate and ensure the quality of one of our discipline's central methods. The release of new methods (Rieger et al.) and a particular dataset (Chan et al.) in this Special Issue also underline the continuing efforts to keep the quality of content analysis as high as possible, often with an increased focus on open science principles (Dienlin et al., 2021; Freiling et al., 2021).

Second, the submissions we received, as well as the six contributions we finally selected for the Special Issue, only partially focused on the five criteria we deemed important. Validity and reliability, the latter mostly defined in the traditional sense of coder agreements, as established quality criteria for content analysis (Krippendorff, 2018; Potter & Levine-Donnerstein, 1999) were addressed most often by far. Reproducibility, robustness, and replicability – which we considered to be new ways of (re)establishing the quality criterion of reliability – were less overtly addressed. Only one submission raised awareness of a quality criterion we had originally not considered explicitly: ethics (Wiedicke, part of this Special Issue). As such, it seems that the field mostly focuses on traditional quality criteria, although these may be partly reinterpreted and approximated in new ways. Examples of such reinterpretations and approximations in the realm of ethics include a dedicated reflection on ownership structures, the important yet in some cases impossible collection of informed consent, and how human coders might be affected by their coding tasks, as well as how they affect subsequent methodological decisions. Overall, this indicates that future work could extend reflections on conceptual clarity and methodological approaches related to traditional quality criteria (validity, reliability), how they may be reestablished (reproducibility, robustness, and replicability), as well as criteria that have recently come into focus, such as ethics.

Relatedly, and third, as pointed out in numerous submissions, human coders are still one of the cornerstones of content analysis in the tradition of communication science. Their coding is capable of making *a*, if not *the*, difference to the quality of any content analysis. Our discipline's experience with this human-in-the-loop method purportedly makes this method a key strength of communication science. However, the use of human coders is currently being challenged severely by technological influence (e.g., Pangakis et al., 2023), which calls for a good middle ground between effort and benefit. Here, we should check for the benefit that human coders can bring to the table (Baden et al., part of this Special Issue), reassure ourselves that potential error sources are held at bay (Oschatz et al., part of this Special Issue), and repeatedly question whether human coders' – admittedly expensive – involvement is still worth the cost (Niemann-Lenz et al., part of this Special Issue).

Fourth, contemporary content analysis is being challenged by a plethora of external adjustments and technological improvements. The integration of technological advancements into communication science seems to be applying the most pressure on the field and its methods, as the many submissions on testing reliability, reproducibility, replicability, and robustness as necessary conditions for validity (Chan et al., Oschatz et al., Rieger et al., all part of this Special Issue) or new forms of data collection, here via crowdcoders (Niemann-Lenz et al., part of this Special Issue), indicate. In turn, fewer challenges to quality criteria seem to come from our field by means of internal debates, which are presented here as conceptual thoughts on valid disagreement (Baden et al., part of this Special Issue). This indicates that change from within the field, both in terms of conceptualizing and testing quality criteria, is an important yet understudied avenue for future work. Our final diagnosis is that more work at these crossroads, primarily between communication and computer science, is also a great chance to extend and even export our expertise. The sophisticated reflection and integration of technological advancements into content analysis (Oschatz et al., part of this Special Issue) is clearly a strength of the methodologically versed part of our field. A change of perspectives could also benefit cross-disciplinary collaboration because content analysis, as a method that can not only stand but also profitably incorporate computer-scientific innovation, might be as interesting to other disciplines as it is to us. As such, while adjusting our quality criteria is important for adapting to but also for resisting external pressure, such conceptual thoughts and methodological advances need to balance inspiration and input from both inside and outside our field.

References

Bachl, M., & Scharkow, M. (2017). Correcting measurement error in content analysis. *Communication Methods and Measures*, *11*(2), 87–104. https://doi.org/10.1080/19312458.2017.1305103

Baden, C., Pipal, C., Schoonvelde, M., & van der Velden, M. A. C. G. (2022). Three gaps in computational text analysis methods for social sciences: A research agenda. *Communication Methods and Measures*, *16*(1), 1–18. https://doi.org/10.1080/19312458.2021.2015574

Casas, A., & Webb Williams, N. (2022). Introduction to the special issue on images as data. *Computational Communication Research*, *4*(1), 1–10. https://doi.org/10.5117/CCr2022.1.000.CASA

Dienlin, T., Johannes, N., Bowman, N. D., Masur, P. K., Engesser, S., Kümpel, A. S., Lukito, J., Bier, L. M., Zhang, R., Johnson, B. K., Huskey, R., Schneider, F. M., Breuer, J., Parry, D. A., Vermeulen, I., Fisher, J. T., Banks, J., Weber, R., Ellis, D. A., … De Vreese, C. (2021). An agenda for open science in communication. *Journal of Communication*, *71*(1), 1–26. https://doi.org/10.1093/joc/jqz052

Domahidi, E., Yang, J., Niemann-Lenz, J., & Reinecke, L. (2019). Outlining the way ahead in computational communication science: An introduction to the IJoC special section on "Computational Methods for Communication Science: Toward a Strategic Roadmap". *International Journal of Communication*, *13*, 3876–3884.

Dutta, M. J., Kaur, S., & Elers, P. (2020). Validity in interpretive methods: Frameworks and innovations. *Annals of the International Communication Association*, *44*(3), 185–200. https://doi.org/10.1080/23808985.2020.1792795

Freiling, I., Krause, N. M., Scheufele, D. A., & Chen, K. (2021). The science of open (communication) science: Toward an evidence-driven understanding of quality criteria in communication research. *Journal of Communication*, *71*(5), 686–714. https://doi.org/10.1093/joc/jqab032

Geise, S., & Waldherr, A. (2021). Computational communication science: Lessons from working group sessions with experts of an emerging research field. In U. Engel, A. Quan-Haase, S. X. Liu, & L. Lyberg (Eds.), *Handbook of computational social science, Volume 1: Theory, case studies and ethics* (pp. 66–82). Routledge. https://doi.org/10.4324/9781003024583

Geiß, S. (2021). Statistical power in content analysis designs: How effect size, sample size and coding accuracy jointly affect hypothesis testing – A Monte Carlo simulation approach. *Computational Communication Research*, *3*(1), 61–89. https://doi.org/10.5117/CCR2021.1.003.GEIS

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Günther, E., & Quandt, T. (2016). Word counts and topic models: Automated text analysis methods for digital journalism research. *Digital Journalism*, *4*(1), 75–88. https://doi.org/10.1080/21670811.2015.1093270

Haim, M. (2021). Gütekriterien und Handlungsempfehlungen für die Entwicklung von Forschungssoftware in der Kommunikations- und Medienwissenschaft [Quality criteria and recommendations for developing research software for communication science and media research]. *Medien & Kommunikationswissenschaft*, *69*(1), 65–79. https://doi.org/10.5771/1615-634X-2021-1-65

Haim, M. (2023). *Computational Communication Science: Eine Einführung* [Computational social science: An introduction]. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-40171-9

Hase, V., Mahl, D., & Schäfer, M. S. (2023). The "computational turn": An "interdisciplinary turn"? A systematic review of text as data approaches in journalism studies. *Online Media and Global Communication*, *2*(1), 122–143. https://doi.org/10.1515/omgc-2023-0003

Helles, R., & Ørmen, J. (2020). Big data and explanation: Reflections on the uses of big data in media and communication research. *European Journal of Communication*, *35*(3), 290–300. https://doi.org/10.1177/0267323120922088

Jünger, J., & Gärtner, C. (2023). *Computational Methods für die Sozial- und Geisteswissenschaften* [Computational methods for the social sciences and humanities]. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-37747-2

Jünger, J., Geise, S., & Hänelt, M. (2022). Unboxing computational social media research from a datahermeneutical perspective: How do scholars address the tension between automation and interpretation? *International Journal of Communication*, *16*, 1482–1505.

Keating, D. M., & Totzkay, D. (2019). We do publish (conceptual) replications (sometimes): Publication trends in communication science, 2007–2016. *Annals of the International Communication Association*, *43*(3), 225–239. https://doi.org/10.1080/23808985.2019.1632218

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 205395171452848. https://doi.org/10.1177/2053951714528481

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage.

Krippendorff, K. (2021). A quadrilogy for (big) data reliabilities. *Communication Methods and Measures*, *15*(3), 165–189. https://doi.org/10.1080/19312458.2020.1861592

Lacy, S., Watson, B. R., Riffe, D., & Lovejoy, J. (2015). Issues and best practices in content analysis. *Journalism & Mass Communication Quarterly*, *92*(4), 791–811. https://doi.org/10.1177/1077699015607338

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*(3), 389–402. https://doi.org/10.1177/2515245918787489

Loosen, W., & Scholl, A. (2012). Theorie und Praxis von Mehrmethodendesigns in der Kommunikationswissenschaft [Theory and practice of mixed-method designs in communication science]. In W. Loosen & A. Scholl (Eds.), *Methodenkombinationen in der Kommunikationswissenschaft: Methodologische Herausforderungen und empirische Praxis* (pp. 9–25). von Halem.

Mahl, D., Von Nordheim, G., & Guenther, L. (2023). Noise pollution: A multi-step approach to assessing the consequences of (not) validating search terms on automated content analyses. *Digital Journalism*, *11*(2), 298–320. https://doi.org/10.1080/21670811.2022.2114920

Nelson, L. K. (2019). To measure meaning in big data, don't give me a map, give me transparency and reproducibility. *Sociological Methodology*, *49*(1), 139–143. https://doi.org/10.1177/0081175019863783

Nelson, L. K., Burk, D., Knudsen, M., & McCall, L. (2021). The future of coding: A comparison of hand-coding and three types of computer-assisted text analysis methods. *Sociological Methods & Research*, *50*(1), 202–237. https://doi.org/10.1177/0049124118769114

Pangakis, N., Wolken, S., & Fasching, N. (2023). *Automated annotation with generative AI requires validation*. arXiv. http://arxiv.org/abs/2306.00176

Pipal, C., Song, H., & Boomgaarden, H. G. (2023). If you have choices, why not choose (and share) all of them? A multiverse approach to understanding news engagement on social media. *Digital Journalism*, *11*(2), 255–275. https://doi.org/10.1080/21670811.2022.2036623

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, *27*(3), 258–284. https://doi.org/10.1080/00909889909365539

RatSWD. (2023). *Erhebung und Nutzung unstrukturierter Daten in den Sozial-, Verhaltens- und Wirtschaftswissenschaften* [Collecting and using unstructured data in the social sciences, behavioural science, and economics]. https://doi.org/10.17620/02671.73

Schatto-Eckrodt, T. (2022). Hidden biases – The effects of unavailable content on Twitter on sampling quality. In J. Jünger, U. Gochermann, C. Peter, & M. Bachl (Eds.), *Grenzen, Probleme und Lösungen bei der Stichprobenziehung* [Limitations, problems, and solutions related to sampling] (pp. 178–195). Herbert von Halem Verlag.

Schoch, D., Chan, C., Wagner, C., & Bleier, A. (2023). *Computational reproducibility in computational social science*. arXiv. http://arxiv.org/abs/2307.01918

Sommer, K., Wettstein, M., Wirth, W., & Matthes, J. (2014). *Automatisierung in der Inhaltsanalyse* [Automation in content analyses]. Herbert von Halem Verlag.

Song, H., Tolochko, P., Eberl, J.-M., Eisele, O., Greussing, E., Heidenreich, T., Lind, F., Galyga, S., & Boomgaarden, H. G. (2020). In validations we trust? The impact of imperfect human annotations as a gold standard on the quality of validation of automated content analysis. *Political Communication*, *37*(4), 550–572. https://doi.org/10.1080/10584609.2020.1723752

TeBlunthuis, N., Hase, V., & Chan, C.-H. (2023). *Misclassification in automated content analysis causes bias in regression. Can we fix it? Yes we can!* arXiv. http://arxiv.org/abs/2307.06483

Theocharis, Y., & Jungherr, A. (2021). Computational social science and the study of political communication. *Political Communication*, *38*(1–2), 1–22. https://doi.org/10.1080/10584609.2020.1833121

van Atteveldt, W., Margolin, D., Shen, C., Trilling, D., & Weber, R. (2019). A roadmap for computational communication research. *Computational Communication Research*, *1*(1), 1–11. https://doi.org/10.5117/CCR2019.1.001.VANA

van Atteveldt, W., Trilling, D., & Arcíla Calderón, C. (2022). *Computational analysis of communication: A practical introduction to the analysis of texts, networks, and images with code examples in Python and R*. Wiley Blackwell.

Walter, N., Cody, M. J., & Ball-Rokeach, S. J. (2018). The ebb and flow of communication research: Seven decades of publication trends and research priorities. *Journal of Communication*, *68*(2), 424–440. https://doi.org/10.1093/joc/jqx015

Wieling, M., Rawee, J., & van Noord, G. (2018). Reproducibility in computational linguistics: Are we willing to share? *Computational Linguistics*, *44*(4), 641–649. https://doi.org/10.1162/coli_a_00330

Wilkerson, J., & Casas, A. (2017). Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science*, *20*(1), 529–544. https://doi.org/10.1146/annurev-polisci-052615-025542