

## Are We Replicating Yet? Reproduction and Replication in Communication Research

Johannes Breuer <sup>1,2</sup>  and Mario Haim <sup>3</sup> 

<sup>1</sup> Department Computational Social Science, GESIS—Leibniz Institute for the Social Sciences, Germany

<sup>2</sup> Research Data & Methods, Center for Advanced Internet Studies, Germany

<sup>3</sup> Department of Media and Communication, LMU Munich, Germany

**Correspondence:** Johannes Breuer ([johannes.breuer@gesis.org](mailto:johannes.breuer@gesis.org))

**Submitted:** 26 March 2024 **Published:** 19 June 2024

**Issue:** This editorial is part of the issue "Reproducibility and Replicability in Communication Research" edited by Johannes Breuer (GESIS—Leibniz Institute for the Social Sciences / Center for Advanced Internet Studies) and Mario Haim (LMU Munich), fully open access at <https://doi.org/10.17645/mac.i429>

### Abstract

The replication crisis has highlighted the importance of reproducibility and replicability in the social and behavioral sciences, including in communication research. While there have been some discussions of and studies on replications in communication research, the extent of this work is significantly lower than in psychology. The key reasons for this limitation are the differences between the disciplines in the topics commonly studied and in the methods and data commonly used in communication research. Communication research often investigates dynamic topics and uses methods (e.g., content analysis) and data types (e.g., media content and social media data) that are not used, or, at least, are much less frequently used, in other fields. These specific characteristics of communication research must be considered and require a more nuanced understanding of reproducibility and replicability. This thematic issue includes commentaries presenting different perspectives, as well as methodological and empirical work investigating the reproducibility and replicability of a wide range of communication research, including surveys, experiments, systematic literature reviews, and studies that involve social media or audio data. The articles in this issue acknowledge the diversity and unique features of communication research and present various ways of improving its reproducibility and replicability, as well as our understanding thereof.

### Keywords

communication research; meta-science; open science; replicability; reproducibility

Scientific progress is based on the premise that the generation of knowledge is cumulative and that science is self-correcting. To ensure that scientific research is cumulative and self-correcting, the reproducibility and replicability of empirical research play key roles. The replicability of research, especially in the social and

behavioral sciences, has become a topic of raised academic but also public interest with the dawn of the so-called “replication crisis” in the early 2010s, sparked by failed attempts at systematically replicating several seminal findings in psychology. Subsequently, similarly low replication rates have also emerged in other disciplines, including more broadly in the social sciences (Camerer et al., 2018).

*Replicability* has multiple definitions. The glossary of the Framework for Open and Reproducible Research Trainin, hence, describes it as “an umbrella term, used differently across fields” (Parsons et al., 2022). A common definition that we also use in this article is that of The Turing Way Community (2022): “A result is replicable when the same analysis performed on different datasets produces qualitatively similar answers.” Notably, there are different types of replications. A common distinction is between a direct or exact replication and a conceptual replication. According to Shrout and Rodgers (2018), “a direct or exact replication is a new study that employs the same procedure, materials, measures, and study population as the original study,” and “a conceptual replication is intentionally different from a direct replication and is designed to assess [the] generalizability, as well as [the] veracity, of a result” (p. 492). Replication differs from *reproduction*, which means that “the same analysis steps performed on the same dataset consistently [produce] the same answer” (The Turing Way Community, 2022). In the wake of the replication crisis, researchers have also started to discuss and assess reproducibility. Although all research should, ideally, at least be reproducible, empirical investigations in the social and behavioral sciences have revealed that substantial parts of published research are not reproducible (see, e.g., Artner et al., 2021) and that the main reasons for this are a lack of shared data and code as well as errors in code execution, and insufficient documentation (Hardwicke et al., 2020; Krähler et al., 2023; Trisovic et al., 2022).

Replicability has also been discussed in communication research, with initial studies on this subject dating back to several decades before the onset of the current replication crisis (Kelly et al., 1979). Following the identification and assessment of the replication crisis in psychology and other disciplines, in 2018, the journal *Communication Studies* published a Special Issue for Replications, collecting a series of nine replication studies from different subfields of communication research (volume 69, issue 3). However, despite these and similar singular efforts, such as the Open Science theme for the 2020 Annual Conference of the International Communication Association or a subsequent issue in the *Journal of Communication* (volume 71, issue 5), what McEwan et al. (2018) wrote in their editorial for the aforementioned special issue is still true: that “direct replications...are few and far between” (p. 236). McEwan et al. also pointed out the following important reasons for this observation: “engaging in replications is often undervalued” and “replication studies can be difficult to publish” (p. 325).

Notably, these challenges are not unique to communication science. However, in communication science, unlike in psychology, the replication crisis has not led to large-scale replication efforts or substantial cultural changes in the ways research is conducted and in the requirements it should meet. This is partly because the replication crisis originated in psychology, but this reason cannot fully explain the differences between psychology and communication science in dealing with questions of replicability. Three aspects that play a major role in this regard are the (a) methods and (b) data types that are commonly used in these fields, as well as the (c) topics that are studied in communication research. Although there are some overlaps in the methods and data types used in psychology and communication (e.g., experiments and self-reports, which are widely used in psychology, are also widely used in communication research), communication researchers also widely employ other methods and data types, such as content analyses, qualitative interviews, media content,

or social media data. This has implications for the nature and relevance of replications not only at the practical level but also at the conceptual level.

Notably, the distinction between reproduction and replication is less clear in communication research than in other fields. For example, numerous communication studies have used X (formerly Twitter) data collected via the platform's application programming interface (API). This API used to be freely and easily usable for academic research, but its terms of service allowed only the sharing of post IDs (not the full texts). Over time, if one were to use shared post IDs to recollect the original data (a process known as rehydration) in order to run the same analyses as in the original study, some posts and accounts might have already been deleted. In addition, API features might have already changed. Thus, the data will likely be different. Does this then constitute a reproduction or a replication?

Apart from this conceptual question of what replicability and reproducibility mean in and for communication research, another important question is when to expect reproducibility and/or replicability in communication research. To give yet another example, in content analyses of news reports at a specific time and on a specific topic, the findings might not be expected to be replicated in empirical efforts conducted on the same topic but several years later or in a different country. This contextual dependency of replicability in communication research starkly contrasts with the case in, for example, large parts of psychology, where successful replications are considered crucial estimates of the robustness and generalizability of research findings.

Aside from conceptual questions on reproducibility and replicability in communication research, there are also empirical questions. These empirical questions are related to (a) the prevalence and degree of reproducibility and replicability and (b) the ways of facilitating and improving reproductions and reproducibility, as well as replications and replicability. However, answering such questions requires conducting actual reproductions and replications, as well as systematic assessments of research practices in the field, and how they can affect reproducibility and replicability. As noted before, the three main factors that set communication research apart from other disciplines and are important to consider when answering conceptual as well as empirical questions related to reproducibility and replicability are: (a) the topics studied in communication research, (b) the methods employed in communication research, and (c) the various types of data that are used by communication scholars.

A key defining feature of many of the topics that are studied in communication research is their dynamic and, in part, also transient nature. Media content, technology, platforms, and usage patterns often undergo rapid transformations. In addition, studies regularly consider topics from current events. Accordingly, unlike studies in psychology that aim to understand the essential or universal mechanisms of human cognition and behavior, for large parts of communication research, cultural and/or temporal differences are expected to strongly affect study outcomes and, thus, to limit replicability.

Although communication research widely uses methods that are also used in other areas of the social and behavioral sciences, such as surveys, experiments, and interviews, one method is originally associated with the discipline: content analysis. Content analyses in communication research often investigate media content and thereby look at timely topics, such as news reporting. Another more recent methodological development in communication research is the increasing use of computational methods, especially within

the growing subfield of computational communication science (see van Atteveldt & Peng, 2018). The use of machine learning algorithms, natural language processing techniques, and network analysis introduces new challenges for reproducibility and replicability, as these methods often involve probabilistic procedures and require documentation of additional steps and decisions, such as data preprocessing or parameter (and, in the case of machine learning, also hyperparameter) selection.

As for the methods, there are some overlaps between communication research and other fields in the social and behavioral sciences with regard to data types. Many studies rely on self-report data, which are typically gathered via surveys, interviews, or questionnaires in experimental studies. However, some data types are much more widely used in communication research, such as media content and social media data. In general, much of communication research draws upon data governed by platform-specific terms of service, proprietary content, or individual-level data that are subject to privacy concerns. This limits the capacity for data sharing due to privacy and copyright regulations (Davidson et al., 2023; van Atteveldt et al., 2020) and, thus can reduce reproducibility.

Overall, the special features of communication research necessitate a more nuanced perspective on reproducibility and replicability. Instead of asking only how reproducible or replicable communication research is, we should also ask what kinds of reproduction and replication are possible and informative, and what is needed to enable or facilitate different kinds of reproduction and replication. For example, in their analysis of communication science studies published between 2007 and 2016, Keating and Totzkay (2019) found that conceptual replications are much more common than direct replications. Considering the topics, methods, and data types that are often used in communication research, conceptual replications may generally be more appropriate than direct or exact replications. Another type of replication that may be particularly suitable for communication science is *prospective replication*, in which researchers plan “a series of replication studies that may occur simultaneously or at different times” (Steiner et al., 2019, p. 281). A helpful clarification question in addition to what type of replication is being pursued is what steps (or parts thereof) should be reproduced or replicated at all: data collection, its processing, the analysis, or the compilation and interpretation of results.

The eight articles included in this thematic issue reflect the breadth and complexity of replicability and reproducibility in communication research and address different conceptual, methodological, and empirical questions in various ways. In particular, this thematic issue combines three replication studies, three methodological articles, and two commentaries.

Knöpfle and Schatto-Eckrodt (2024), in their combined reproduction and replication study, empirically assess the challenges in working with social media data, in this case, with data from X. Although the findings from the original study could largely be reproduced using the same data, in the replication attempt, only slightly more than half of the posts could be recollected, which led to substantial differences in the results. Dubèl et al. (2024) also replicate a previous study that found that viewers are more aroused by negative than positive news, by conducting a laboratory study that combined physiological measures with self-reports. They also extend the result of the previous study by repeating the study in another country and using additional measures.

Adopting a methodological perspective, Lukito et al. (2024) present their study on the implications of tool choice and preprocessing of audio data for the reproducibility and replicability of studies that use this data

type. Although they found that the tools they tested provided accurate automated transcriptions, they note subtle yet significant differences between tools that could also impact reproducibility and replicability. Rogge et al. (2024) propose a standardized sampling method as a way of ensuring reproducibility and replicability for systematic literature reviews. Their method represents a structured multistage approach that can complement and extend existing guidelines for systematic literature reviews. The article of Vermeulen et al. (2024) offers a conceptual and practical exploration of direct replication in experimental communication science. They argue that replication studies in communication research almost always require the adaptation of at least some parts of the original design, extend existing replication typologies by adding the dimension of the motivation behind a replication study, and provide recommendations for replicators.

Xu and Zhang (2024) replicated and extended an earlier survey study on data-sharing practices among psychologists with a sample of Chinese communication scholars. Building on the theory of planned behavior, they find that various factors, including perceived risks and benefits, subjective norms, and pressure from journals, influence attitudes toward data sharing. Based on their findings, they present practical suggestions for improving research practices that can facilitate reproduction and replication in communication research. In his commentary, Ivory (2024) illustrates the urgent implications of the replication crisis for communication research. Acknowledging different perspectives on issues related to reproducibility and replicability and potential solutions, his discussion focuses on the dimensions of responsibility to the public, stewardship of resources, and membership in a community of scholars. Finally, in another commentary, Bowman (2024) stresses the continued need for replications in communication research. He reflects on key issues and recent developments and discusses replication as a key element of postpositivist approaches.

We find the collection of articles in this thematic issue, taken together, not only insightful but also highly representative of the breadth of perspectives and challenges pertaining to reproducibility and replicability in communication research. We believe that the different contributions in this issue can help in arriving at a better understanding of the nature and relevance of reproducibility and replicability in communication research, as well as of potential challenges and ways to address those.

### **Acknowledgments**

First and foremost, we thank the authors of the articles in this thematic issue for their contributions. We also thank the reviewers of this thematic issue. Their expert feedback on the articles has helped significantly in ensuring and improving the overall quality of the issue. We are also grateful to the members of the projects in the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) priority program META-REP: A Meta-Scientific Programme to Analyse and Optimise Replicability in the Behavioral, Social, and Cognitive Sciences, with whom we had many fruitful and inspiring exchanges about reproducibility and replicability and their meaning and assessment, as well as ways of improving them. Finally, we extend our appreciation to the editorial team at *Media and Communication* for their support in the entire process of planning and publishing this thematic issue.

### **Funding**

The editing of this thematic issue is part of the work in the project What Defines and Affects Replicability in Computational Communication Science? (project number 464291459) funded by the DFG under the META-REP priority program (project number 441890184).

## Conflict of Interests

The authors declare no conflict of interests.

## References

- Artner, R., Verliefde, T., Steegen, S., Gomes, S., Traets, F., Tuerlinckx, F., & Vanpaemel, W. (2021). The reproducibility of statistical results in psychological research: An investigation using unpublished raw data. *Psychological Methods, 26*(5), 527–546. <https://doi.org/10.1037/met0000365>
- Bowman, N. D. (2024). On the continued need for replication in media and communication research. *Media and Communication, 12*, Article 7935.
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., . . . Wu, H. (2018). Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour, 2*(9), 637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- Davidson, B. I., Wischerath, D., Racek, D., Parry, D. A., Godwin, E., Hinds, J., Van Der Linden, D., Roscoe, J. F., Ayravainen, L., & Cork, A. G. (2023). Platform-controlled social media APIs threaten open science. *Nature Human Behaviour, 7*, 2054–2057. <https://doi.org/10.1038/s41562-023-01750-2>
- Dubèl, R., Schumacher, G., Homan, M. D., Peterson, D., & Bakker, B. N. (2024). Replicating and extending Soroka, Fournier, and Nir: Negative news increases arousal and negative affect. *Media and Communication, 12*, Article 7807.
- Hardwicke, T. E., Wallach, J. D., Kidwell, M. C., Bendixen, T., Crüwell, S., & Ioannidis, J. P. A. (2020). An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014–2017). *Royal Society Open Science, 7*(2), Article 190806. <https://doi.org/10.1098/rsos.190806>
- Ivory, J. D. (2024). Remembering reasons for reform: A more replicable and reproducible communication literature without the rancor. *Media and Communication, 12*, Article 7852.
- 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>
- Kelly, C. W., Chase, L. J., & Tucker, R. K. (1979). Replication in experimental communication research: An analysis. *Human Communication Research, 5*(4), 338–342. <https://doi.org/10.1111/j.1468-2958.1979.tb00646.x>
- Knöpfle, P., & Schatto-Eckrodt, T. (2024). The challenges of replicating volatile platform-data studies: Replicating Schatto-Eckrodt et al. (2020). *Media and Communication, 12*, Article 7789.
- Krähmer, D., Schächtele, L., & Schneck, A. (2023). Care to share? Experimental evidence on code sharing behavior in the social sciences. *PLoS ONE, 18*(8), Article e0289380. <https://doi.org/10.1371/journal.pone.0289380>
- Lukito, J., Greenfield, J., Yang, Y., Dalhke, R., Brown, M. A., Lewis, R., & Chen, B. (2024). Audio-as-data tools: Replicating computational data processing. *Media and Communication, 12*, Article 7851.
- McEwan, B., Carpenter, C. J., & Westerman, D. (2018). On replication in communication science. *Communication Studies, 69*(3), 235–241. <https://doi.org/10.1080/10510974.2018.1464938>
- Parsons, S., Azevedo, F., Elsherif, M. M., Guay, S., Shahim, O. N., Govaart, G. H., Norris, E., O'Mahony, A., Parker, A. J., Todorovic, A., Pennington, C. R., Garcia-Pelegrin, E., Lazić, A., Robertson, O., Middleton, S. L., Valentini, B., McCuaig, J., Baker, B. J., Collins, E. . . . Aczel, B. (2022). A community-sourced glossary of open scholarship terms. *Nature Human Behaviour, 6*(3), 312–318. <https://doi.org/10.1038/s41562-021-01269-4>



- Rogge, A., Anter, L., Kunze, D., Pomsel, K., & Willenbrock, G. (2024). Standardized sampling for systematic literature reviews (STAMP method): Ensuring reproducibility and replicability. *Media and Communication*, 12, Article 7836.
- Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, 69(1), 487–510. <https://doi.org/10.1146/annurev-psych-122216-011845>
- Steiner, P. M., Wong, V. C., & Anglin, K. (2019). A causal replication framework for designing and assessing replication efforts. *Zeitschrift Für Psychologie*, 227(4), 280–292. <https://doi.org/10.1027/2151-2604/a000385>
- The Turing Way Community. (2022). *The Turing way: A handbook for reproducible, ethical and collaborative research*. Zenodo. <https://doi.org/10.5281/zenodo.3233853>
- Trisovic, A., Lau, M. K., Pasquier, T., & Crosas, M. (2022). A large-scale study on research code quality and execution. *Scientific Data*, 9, Article 60. <https://doi.org/10.1038/s41597-022-01143-6>
- van Atteveldt, W., Althaus, S., & Wessler, H. (2020). The trouble with sharing your privates: Pursuing ethical open science and collaborative research across national jurisdictions using sensitive data. *Political Communication*, 38(1/2), 192–198. <https://doi.org/10.1080/10584609.2020.1744780>
- van Atteveldt, W., & Peng, T.-Q. (2018). When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science. *Communication Methods and Measures*, 12(2/3), 81–92. <https://doi.org/10.1080/19312458.2018.1458084>
- Vermeulen, I., Masur, P. K., Beukeboom, C. J., & Johnson, B. J. (2024). Direct replication in experimental communication science: A conceptual and practical exploration. *Media and Communication*, 12, Article 7971.
- Xu, J., & Zhang, R. (2024). Attitudinal, normative, and resource factors affecting communication scholars' data sharing: A replication study. *Media and Communication*, 12, Article 7666.

## About the Authors



**Johannes Breuer** (PhD) is a senior researcher and leader of the team Digital Society Observatory in the Department of Computational Social Science at GESIS—Leibniz Institute for the Social Sciences in Cologne, Germany, and the team Research Data and Methods at the Center for Advanced Internet Studies (CAIS) in Bochum, Germany. His research interests include the use and effects of digital media, digital trace data and computational methods, as well as open science and meta-science. More information: <https://www.johannesbreuer.com>



**Mario Haim** (PhD) is full professor for Communication Science with a focus on Computational Communication Research at the Department of Media and Communication at LMU Munich, Germany. His research interests circle around algorithmic influences, such as in political communication, journalism, health communication, or media use, as well as on (computational) methods and meta-science. More information: <https://haim.it>. Photo by Lena Fleischer.